DOCUMENT RESUME

ED 143 545 SE 023 029

AUTHOR	Haag, Vincent H.
TITLE	Studies in Mathematics, Volume III. Structure of Elementary Algebra. Revised Edition.
INSTITUTION	Stanford Univ., Calif. School Mathematics Study Group.
SPONS AGENCY	National Science Foundation, Washington, D.C.
PUB DATE	61
NOTE	233p.; For related documents, see SE 023 028-041
EDRS PRICE	MF-$0.83 HC-$12.71 Plus Postage.
DESCRIPTORS	*Algebra; Inservice Education; *Instructional Materials; Resource Materials; Secondary Grades; *Secondary School Mathematics; Teacher Education; *Teaching Guides; Textbooks
IDENTIFIERS	*School Mathematics Study Group

ABSTRACT

These materials are intended to explain the approach adopted by the writers of the SMSG textbook, First Course in Algebra. This book is not a ninth-grade textbook or teacher's commentary. Many of the ideas presented are too difficult for most beginning students, but they are ideas which the author believes teachers should master. It is assumed that the teacher already masters these skills. Chapters included are: (1) Historical Background; (2) Language; (3) Structure of the Real Number System; (4) Sub-Systems of the Real Numbers; (5) Completeness of the Real Number System; and (6) Functions. The appendices include materials on infinite decimals, complex numbers, algebraic numbers, and answers to exercises. (RH)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED143545

SCHOOL
MATHEMATICS
STUDY GROUP

STUDIES IN MATHEMATICS
VOLUME III

Structure of Elementary Algebra

(revised edition)

By VINCENT H. HAAG

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

SMSG

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

2

SE 023 029

SCHOOL
MATHEMATICS
STUDY GROUP

STUDIES IN MATHEMATICS
VOLUME III

Structure of Elementary Algebra

(revised edition)

By VINCENT H. HAAG, Franklin and Marshall College

Written for the SCHOOL MATHEMATICS STUDY GROUP
Under a grant from the NATIONAL SCIENCE FOUNDATION

Copyright 1961 by Yale University
PRINTED IN THE U.S.A.

CONTENTS

Preface 1

Chapter

1. HISTORICAL BACKGROUND 1.1
    1. Classical Algebra 1.1
    2. Transition to Modern Algebra 1.5
    3. Structure 1.6
    4. Teaching of Algebra 1.19
    5. A Program for Elementary Algebra 1.21

2. LANGUAGE 2.1
    1. Sets 2.1
    2. Sentences 2.10
    3. Logic 2.21

3. STRUCTURE OF THE REAL NUMBER SYSTEM 3.1
    1. Axioms for an Ordered Field 3.1
    2. Field Axioms 3.2
    3. Order Axioms 3.16
    4. Development of the Real Numbers in SMSG-F . 3.27

4. SUB-SYSTEMS OF THE REAL NUMBERS 4.1
    1. The Natural Numbers 4.1
    2. The Integers 4.10
    3. The Rational Numbers 4.17

5. COMPLETENESS OF THE REAL NUMBER SYSTEM 5.1
    1. The Completeness Axiom 5.1
    2. The Existence of $\sqrt{2}$ in R 5.8
    3. Completeness of the Set of Reals 5.14

6. FUNCTIONS 6.1
    1. Variables 6.1
    2. Algebraic Expressions 6.2
    3. Open Sentences 6.8
    4. Functions 6.18

Appendix

A. INFINITE DECIMALS A.1
    1. Decimal Representations of Real Numbers . . A.1
    2. R is Not Countable A.6

B. COMPLEX NUMBERS B.1

C. ALGEBRAIC NUMBERS C.1

Answers to Exercises . . . . following page C.6

Preface

Part of the activity of the School Mathematics Study Group (SMSG) is devoted to the preparation of experimental mathematics textbooks for secondary schools. This set of notes, the third in a series, is intended to explain the approach adopted by the writers of the textbook First Course in Algebra (SMSG-F). It is expected that these notes may be used in the Summer of 1960 by teachers who are studying SMSG-F and who are familiarizing themselves with new teaching materials in algebra.

It must be understood that this book is not a ninth grade textbook or a teacher's commentary. The ideas presented are far too difficult for most beginning students, but these are ideas which we believe teachers should master. The terminology and notation are the same as, but the topics do not closely parallel, those of the SMSG-F textbook; hence, it is unsuitable as a manual. The notes delve into the foundations of algebra, the structural properties of elementary algebraic systems, but are not concerned with the routine skills and manipulative aspects of algebra. It is assumed that the teacher is already a master of these skills.

In short, this book is not designed to explain how one should teach the SMSG-F materials but rather to explain what is meant by "modern algebra", what concepts underly the SMSG-F materials, and what is the spirit of the materials. It is believed that a teacher who understands these underlying concepts will be able to use the textbook effectively and to stimulate mathematical curiosity in his students.

In particular, it should be pointed out that these notes are not intended for an abstract algebra course, since they are geared directly to SMSG-F.

The instructor of a summer course for teachers will probably find that Chapters 2, 3, 4 and 6 form the heart of the study. Chapter 1 can be read quickly as an introduction, but Chapter 5 may be rough sledding. Certain readers may want to read only the summary results of Chapter 5, and others may be challenged to follow the proofs in detail and try their hands at the problems. In Chapter 2 we discuss the questions, "What is a proof?" and "Why bother with proofs in algebra?"

It is in Chapter 3 that the teacher comes to grips with the structure of algebra. Here we postulate a system called an ordered field (the set of real numbers is the most familiar model of this system) and study its properties. Theorems marked with a star * are left to the reader for proof; these are essential to the development and should be regarded as strongly recommended exercises. Then in Chapter 4 the various subfields of the real numbers are examined, putting them in perspective with each other and with the reals. In Chapter 6 we summarize the relations, operations and expressions of algebra by unifying them under the concept of function.

The outline for these notes was drawn up and the resulting manuscript was read by an advisory committee consisting of:

C.W. Curtis, University of Wisconsin,

B.J. Pettis, University of North Carolina, H.O. Pollak, Bell Telephone Laboratories,

C.E. Rickart, Yale University.

The author is indebted to this committee for its many valuable suggestions.

The instructor should feel free to consider topics in any order and to supplement with ideas and problems at will. No attempt is made (or intended) to freeze the approach to elementary algebra in the present mold. There is a healthy debate going on as to the best way to design a first course in algebra. It is hoped that some readers will learn enough about the issues so that they can enter the debate on one side or the other. With the goal of superlative teaching and learning of algebra in the schools, the reader is invited to convey his comments and criticisms to

School Mathematics Study Group
Drawer 2502 A Yale Station
New Haven, Connecticut

Chapter 1

HISTORICAL BACKGROUND

1. Classical Algebra. What is the genesis of the school book algebra as it is now taught in our schools? How does it differ from "modern algebra" of 20th century mathematics? In what respects have the two parted company, and why? We might find answers to these questions in a short historical sketch.

Somewhere in the haze of pre-history a passage was made from the concrete to the abstract. The idea of "two" as a characteristic of each pair of objects must have taxed the primitive mind, just as it eludes the mind of the very young child.

Number sense came slowly. As systems of notation were invented, the meaning of number became clearer. Histories of mathematics* trace this development along with a growing sense of spatial form, through the early emergence of arithmetic and geometry in

---

* CF. D.E. Smith, History of Mathematics, Ginn, 1923; E.T. Bell, The Development of Mathematics, McGraw-Hill, 1940; D.J. Struik, A Concise History of Mathematics, Dover, 1948; or other histories by Cantor, Hofmann, Eves, etc.

Babylon and Egypt, to the amazing Greek period of deductive reasoning in geometry and the invention of some algebraic symbolism by the Hindus.

Somewhere in this fumbling for satisfaction of man's curiosity about numbers, another step was taken. Certain symbols, or "numerals", had already been used to stand for certain numbers, such as the Persian notation of ۱ for "one", ۲ for "two", ۳ for "three", etc. If a number, not known, was observed to have certain relationships to known numbers, this fact could be described by representing the unknown number by still another kind of symbol. Consider the problem: some number symbolized by α is such that α multiplied by α and then diminished by ۱ yields ۳ ; what number is α ? The solution of such a problem, the bringing together of the known and unknown parts, was called *al-jebr* in Arabic, and *algebra* in medieval Latin, meaning "reunion of parts."

Centuries later, European mathematicians would write this problem in the form of an equation to be solved:

$$x^2 - 1 = 3.$$

But having found one root, 2, they were momentarily satisfied. It was not until after 1600 A.D. that -2 was reluctantly given the status of a number and admitted as another root.

The more general quadratic equation

$$x^2 + bx + c = 0,$$

where b and c are any rational numbers, was "solved" in 1519 by Ferro of the University of Bologna (and as early as the 9th century by the Arabs, it is believed). The cubic and quartic equations were "solved" before 1545 by Tartaglia, Cardan, and Ferrari (leading to a dispute as to whether or not Cardan filched the procedure from Tartaglia). None of these solutions inferred any understanding of negative or imaginary roots.

We must stop here and examine the meaning of "solution." Among the possible numerical values of the variable x, one which makes the equation a true sentence is called a root of the equation. "Solving" an equation then means finding the set of all roots of the equation. To mathematicians of the 16th century there were two distinct meanings.

(1) Approximate solution. Given the numerical coefficients of the equation, construct a numerical approximation to a root of the equation either by geometric construction or by successive refinements of an original estimate of the root*. The Chinese are believed to have effected approximate solutions as early as the 13th century. Such a solution is always possible for a polynomial equation of any degree, if it has a root, and this fact is of immense importance to applications.

---

* For example, we estimate that a root of $x^3 + x - 1 = 0$ is approximately .7 because $(.7)^3 + (.7) - 1 = .043$ and $(.6)^3 + (.6) - 1 = -.184$. This suggests the refined estimate .68 because $(.68)^3 + (.68) - 1 = -.006$ and $(.69)^3 + (.69) - 1 = .019$, etc. Horner's method, Newton's method, and others, are examples of such approximate solutions.

(2) Solution by radicals. Given the rational coefficients of the equation, say, a,b,c,...,k, and given the set of operations +, ×, -, ÷, and extraction of roots, construct by means of a finite number of these operations all expressions in the coefficients which satisfy the equation. Thus, the quadratic

$$x^2 + bx + c = 0$$

can be solved by radicals, because the expression in b,c,

$$\frac{-b \pm \sqrt{b^2 - 4c}}{2} ,$$

satisfies the equation and is constructed by means of only a finite number of elementary operations on b and c.

It was in the sense of solution by radicals that Tartaglia and Ferrari solved the cubic and the quartic, respectively**. Their solutions were sheer monuments of ingenuity, for there were no underlying principles on which they proceeded. Immediately there began a flurry of effort to solve the quintic. It was perhaps natural (although incorrect) to suppose that a solution of the quintic by radicals awaited the man clever enough to discover it. But ingenuity could not prevail, and throughout the 18th century the problem of the quintic remained unsolved.

** The cubic $x^3 + ax^2 + bx + c = 0$ is solved by substituting $x = y - (a/3)$, yielding the reduced cubic $y^3 + py + q = 0$. Then the substitution $y = z - (p/3z)$ reduces this to $z^6 + qz^3 - (p^3/27) = 0$, which is a quadratic in z^3. Thus, the cubic is reduced to a quadratic. The quartic is solved by reduction to a cubic. See a text in theory of equations, such as L.E. Dickson, First Course in Theory of Equations.

In the meanwhile some mathematicians were beginning to examine the methods of constructing solutions. For example, in 1770 Lagrange showed that the method used to construct solutions of the quartic by radicals could not be extended to the solution of the quintic. There the matter stood until 1824.

In summary, at the beginning of the 19th century algebra consisted of a set of rules and devices for performing formal operations on real numbers and symbols representing real numbers (manipulations of algebraic expressions), solutions by radicals of polynomial equations up to the fourth degree, and approximate solutions of polynomials of any degree. This we think of as classical algebra; it is the algebra presented today in traditional elementary textbooks.

2. Transition to Modern Algebra. There was a growing suspicion at the beginning of the 19th century that the quintic may not be solvable by radicals. Possibly there was something inherent in the structure of the real numbers which made the quintic essentially different from the quartic. Then in 1824 a Norwegian, Niels Hendrik Abel, proved that it is impossible to solve the general quintic by radicals. And in 1830 the Frenchman, Evariste Galois, discovered necessary and sufficient conditions for the solution by radicals of *any* polynomial equation. At first thought, one might be tempted to understand the previous sentences to mean that every possible device for solving the quintic was tried and found lacking

--hence, a solution is impossible. But since there are a finite number of operations allowed, after n such operations there is always an (n+1)st operation possible. Thus, Galois and Abel could not have exhausted every combination of operations.

Instead, they searched for properties (characteristics, descriptions) of the equations which isolate the nature of the equations independent of the specific numerical coefficients. Galois set for himself the general problem of determining when an arbitrary polynomial equation with rational coefficients could be solved by radicals. He gave a complete solution to this problem as an application of a general theory of "groups". (We shall examine what is meant by a "group" of elements in the next section.) Differences in groups were found to depend on the relations among their elements rather than on the elements themselves. Thus, Galois could then make statements about roots of equations by noting properties of certain corresponding groups.

At this point we shall not try to explain how Galois constructed groups corresponding to equations. The point is that a break had been made from the classical algebra. It was finally realized that more could be learned about the nature of algebra by studying the structures of mathematical systems, such as groups, than by trying more manipulations on more symbols with more operations.

3. Structure. In the preceding paragraph we used several new words when we indicated that a "group" is an example of a "mathematical system" whose "structure" needs to be studied. Before defining these words, let us gain some preparatory

experience with the ideas involved.

Consider the set of four integers {0,1,2,3}. (Notice that the word "set" has the usual meaning of collection, class or aggregate of elements; we usually indicate a set of elements by enclosing the elements in brackets.) Select any element of this set, say 3, and then again select any element, say 2. Now let us associate with this *ordered pair* of elements a number in the following way. Determine the sum of 3 and 2, divide the sum by four, and find the remainder. Let us indicate this result by writing

$$3 \oplus 2 = 1.$$

(We use "=" to mean that the symbols "$3 \oplus 2$" and "1" represent the same element.) In the same way,

$$1 \oplus 2 = 3, \quad 2 \oplus 2 = 0, \quad 3 \oplus 1 = 0, \quad \text{etc.}$$

Here we have defined a binary operation $\oplus$ on ordered pairs of elements of the set {0,1,2,3}; we say

> a *binary operation* on a set is a rule whereby to each ordered pair of elements of the set there corresponds exactly one element.

For the above example we can show all the results of the operation $\oplus$ in tabular form.

$\oplus$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

,

where the first element is chosen from the left column, the second element from the top row, and the result of the operation in the corresponding row and column of the table.

If the operation associates with each single element of a set exactly one element, we say the operation is *unary*. Familiar examples of binary operations are ordinary addition and multiplication; some common unary operations are squaring and doubling.

The above set $\{0,1,2,3\}$ with the operation $\oplus$ is an example of a mathematical system.

> A *mathematical system* consists of a set of elements and one or more operations on the elements.

If we denote the set $\{0,1,2,3\}$ by the letter T, then the above mathematical system may be denoted by $(T, \oplus)$.

Let us examine some of the properties of $(T, \oplus)$. (By *property* we mean a relationship among elements and operations which is true for all the elements.)

1) The first thing we notice is that every entry in the table is an element of T. More precisely, if a,b are any elements of T, then $a \oplus b$ is an element of T. We say in general that

   > a set S is *closed under a binary operation* $*$ if for any elements x,y in S, $x*y$ is an element of S.

2) We also notice a symmetry in the table. This is the result of a property of $\oplus$ that can be described as follows: if a,b are any elements of T, then $a \oplus b = b \oplus a$. We say that an operation having this property is *commutative*.

3) The reader should verify that the binary operation $\oplus$ is also associative; that is, for any elements a,b,c in T,

$$a \oplus (b \oplus c) = (a \oplus b) \oplus c.$$

4) Among the elements of T we call the element 0 an identity element for $\oplus$ because $0 \oplus a = a \oplus 0 = a$ for any element a in T. That is, any element of T is left unchanged when operating with the identity. In general,

> an element i of a set S is an identity for the operation * if $x*i = i*x = x$ for every x in S.

5) Inspection of the table shows that each row and each column contains the identity element exactly once. This follows from the fact that each element of T has an inverse under $\oplus$: to each element a in T there corresponds an element b in T such that $a \oplus b = b \oplus a = 0$. In general,

> if i is the identity for the operation * in a set S, then x and y are inverses under * if
>
> $$x*y = y*x = i.$$

Of course, in a system the elements may be any objects whatsoever and the operations completely arbitrary. Algebra is concerned not with the elements or the symbols for the elements of a system; it is interested in the structure of the system as described by the basic properties which its operations possess.

The elements of T are quite specific in our minds: the integers 0,1,2,3. The operation $\oplus$ is also specific because it involves the familiar operations of addition and dividing by 4.

The resulting system (T, $\oplus$) therefore has properties which are not surprising to us; in fact, we are led to these properties by our intuitive notions about integers and about addition and dividing by 4. To avoid the prejudices of intuition let us try to forget the meanings of 0,1,2,3 and $\oplus$. Instead, let us write, respectively, s,m,t,r and *. Then the table looks like this.

*	s	m	t	r
s	s	m	t	r
m	m	t	r	s
t	t	r	s	m
r	r	s	m	t

The resulting system is *abstract* in the sense that the symbols are undefined and the operation * has no meaning other than that given by the table. Let us call this the abstract (S,*) system.

Relieved of our pre-set ideas about integers, we might be able to discover hidden properties of the (S,*) system that are inherent in the table. Then since the table for (T, $\oplus$) has exactly the same form as the table for (S,*), that is, the systems have the same *structure*, what we discover about (S,*) must also be true for (T, $\oplus$). In this way we may discover properties of a familiar system that we never suspected through the device of studying the structure of an abstract system.

We say that the specific system (T, $\oplus$) is a *model* of the abstract system (S,*). Many other models can be formed merely by giving other specific meanings to s,m,t,r and *. Thus, a system may admit of many different models, each with the same structure as the system which it is modeling. But two systems are different only if their structures are different.

This connection between systems, models, and structure can be further illustrated by more elementary examples. It is easy to invent an abstract system by choosing any set of elements, writing out a double entry table and filling in the cells arbitrarily with elements. Then an operation $\circ$ is defined from the table by letting $x \circ y$ be the element in row x and column y.

But the systems that people construct are usually chosen because some specific models of the system have appeared elsewhere in mathematics, in physics, or in some other field. As an example, let a set have two elements: a, "the action of reversing an electric switch," and b, "the action of not reversing the switch." If the operation $\circ$ is defined so that $x \circ y$ is the action which has the same result as performing action x and then performing action y, then the system is described by the table:

o	a	b
a	b	a
b	a	b

Note that $a \circ a$ means "reverse the switch and then reverse it again," which has the same result as b, "not reversing the switch." In the same way, $a \circ b = a$, etc. This switching system has the same properties that we observed before. For example, in this system the binary operation $\circ$ is commutative; that is,

$$x \circ y = y \circ x$$

for any replacement of x and y by a or b. The reader should decide whether the operation is also associative, that is, whether

$$x \circ (y \circ z) = (x \circ y) \circ z.$$

It is interesting to note that b is an identity element for $\circ$ in this system, because

$$b \circ a = a \circ b = a \quad \text{and} \quad b \circ b = b.$$

Does every element of this system have an inverse under $\circ$? What is the inverse of a under $\circ$?

Another simple system can be constructed out of the arithmetic of odd and even integers. Let E, O be the elements of the set, and let + symbolize the operation defined by the table:

+	E	O
E	E	O
O	O	E

(An even integer added to an even integer yields an even integer, etc. Here the operation + is not quite the same as the usual addition of numbers. Nor are E and O numbers themselves; they are symbols for classes of numbers. The equation $E + O = O$, for example, means that the sum of any even and any odd numbers is some odd number.) It is left for the reader to verify that the properties of this system are exactly the same as those of the system of switching actions.

These two systems really are two different models of one abstract system consisting of a set of two elements and one binary operation defined by:

*	e	f
e	f	e
f	e	f

where e and f are arbitrary symbols for the elements and * for the binary operation. What is of algebraic concern here is not that the elements and operation can be given various physical or numerical interpretations (although this is interesting), but rather that the three tables have identical structure. Hence, whatever properties we discover in the abstract e,f- system are guaranteed to hold for any model of the system.

The abstract systems we used as examples were selected to illustrate the type of abstract system called a group.

> Given any set S of elements and one binary operation * on elements of S, the system (S,*) is a group if it has the following properties:
>
> (1) For any elements x and y in S, x*y is in S. (S is closed under *.)
>
> (2) For any elements x, y and z in S,
>
> $$x*(y*z) = (x*y)*z.$$
>
> (* is associative.)
>
> (3) There is an element i in S such that
>
> $$x*i = i*x = x$$
>
> for every x in S. (There is an identity for *.)
>
> (4) Corresponding to each element x in S there is an element x' in S such that
>
> $$x*x' = x'*x = i.$$
>
> (Each element has an inverse under *.)

Thus, the first system we studied, $(T, \oplus)$, is a group; it also has the additional property of commutativity and is called a commutative, or abelian, group. The system consisting of the set of all integers and the binary operation of addition is also an abelian group, as the reader should verify. On the other hand, the system consisting of the set of all integers and the operation of multiplication is not a group; it lacks one of the required properties. (Which one?)

Later we shall encounter systems with two binary operations, such as a ring (see Problem 8) and a field (see Chapter 3). The study of the properties of a field is central to the understanding of elementary algebra.

In a rough sort of speaking we can say that before the 19th century, mathematicians were concerned with finding specific entries in tables (studying models, particularly numerical models). This activity took the form of operations on complex combinations of elements, usually algebraic expressions. Since that time most significant discoveries have been made in algebra by studying the structures of abstract systems without regard for the models suggested. It is almost paradoxical that the latter approach turns out to be the most "practical" in every sense of the word.

1.15

Exercises

1. Consider the set of elements {E, O} and the binary operations +, × defined by:

+	E	O
E	E	O
O	O	E

×	E	O
E	E	E
O	E	O

Show that the operation × is commutative. Is there an identity element for × in this set? Determine whether × is distributive through +, that is, whether

$$x \times (y + z) = (x \times y) + (x \times z)$$

for any replacements of E or O for x, y, z. Is + distributive through × ?

2. Consider the system consisting of the set of elements {r, s, t} and the binary operations o, • defined by the tables:

o	r	s	t
r	r	s	t
s	s	t	r
t	t	r	s

•	r	s	t
r	t	s	r
s	r	t	s
t	s	r	t

Is the set closed under o? Under •? Is the operation o commutative? Is • commutative? Is there an identity for o? For •? Is o distributive through •? Is • distributive through o? Does every element have an inverse under o?

3. Let the elements of a set of actions be:

   A: rotating an equilateral triangle 120° clockwise about its center in its plane

   B: rotating the equilateral triangle 240° clockwise about its center in its plane

   C: not rotating the triangle.

   Let $x \circ y$ be the action which has the same result as first performing action x and then performing action y. Construct a table showing all results of the operation. Does this set and this binary operation form an algebraic system? If so, is the set closed under $\circ$? Is the operation $\circ$ commutative? Associative? Is there an identity element for the operation? Does every element have an inverse under $\circ$? Is this system a group? Is it a commutative group?

4. Consider a set of four actions consisting of the four rotations of a square analogous to those of a triangle as described in Problem 3. Is the resulting system a group? A commutative group?

5. Consider the set $\{1,2,3,4\}$ and the operation $\odot$ defined as follows: for any elements a,b in the set, $a \odot b$ is the remainder upon dividing the product of a and b by 5. For example, $3 \odot 4 = 2$, $4 \odot 4 = 1$, etc. Is the resulting system a group?

6. Determine whether the set {a,b,c,d} and the operation ‡, as defined by the following table, is a group.

‡	a	b	c	d
a	b	d	a	c
b	d	c	b	a
c	a	b	c	d
d	c	a	d	b

If not, what properties are lacking?

7. Determine whether the set {r,s,u,v} and the operation ++, as defined by the following table, is a group.

++	r	s	u	v
r	r	s	u	v
s	s	r	v	u
u	u	v	r	r
v	v	u	s	s

If not, what properties are lacking?

8. If (S,•) is a commutative group and if o is another binary operation on elements of S such that o is associative and o is distributive through • (see Problem 1), then the system (S,•,o) is called a ring. Is the set {E,O} and the operations +, ×, as defined in Problem 1, a ring?

1.18

9. Consider the set I of all integers and the operations of ordinary addition + and multiplication ×. Is the system $(S,+,\times)$ a ring? In this system, is there an identity for ×? Does every element of I have an inverse under ×? Is it a commutative ring (that is, is × commutative)?

10. Consider the set $\{0,1,2,3\}$ and the operations $\oplus$, $\odot$ defined on this set as follows:

    $a \oplus b$ is the remainder upon dividing $a + b$ by 4,

    $a \odot b$ is the remainder upon dividing ab by 4.

    Decide whether this system is a ring. A commutative ring.

11. Prove that the identity element of a group $(S,\bullet)$ is unique. (Assume two different identities for $\bullet$, say i and i', and show that this assumption leads to a contradiction.)

12. Prove that for a given element x of a group $(S,\bullet)$ there is a unique inverse under $\bullet$.

13. (a) Do the even integers form a group with respect to addition?
    (b) Do the odd integers form a group with respect to addition?
    (c) Do the integers of the form $5k$, where k is an integer, form a group with respect to addition?
    (d) Let $a \bullet b = a - b$, where a,b are integers. Do the integers form a group with respect to $\bullet$?
    (e) Let $a \odot b$ be the remainder upon dividing ab by 4, where a,b are in the set $\{1,2,3\}$. Does this set form a group with respect to $\odot$?

4. Teaching of Algebra. It is unfortunately true that the description of a specific model of an abstract system brings little understanding of the system. Young students learn many facts about real numbers -- this is an important part of their educations -- but these facts in themselves bring little understanding of the real number system.

Even though the break-through to modern algebra came a hundred years ago, for most school children the word "algebra" still means a collection of isolated tricks -- to each situation a device for handling it. The standard textbook is full of symptoms of this: There are, for example, boxes which emphasize the "how to", or hands pointing to the rule that must be remembered. Some students see for themselves a bit of the structure underlying these tricks. And others enjoy the sheer fun of getting the right "answers" to the manipulations. But for the vast majority it is a matter of memorizing a set of symbolic commands, often in the form of "four step" methods or "rules of signs", etc.

Fortunately, the algebra currently taught in the schools is for the most part mathematically important. The student does need the skills of "symbol pushing" for his later mathematical studies. Hence, it should not be the aim of a new program to change this content drastically. The point, however, is that every bit of manipulation which we teach, and which the student must be able to do, is valid for a reason. There is a mathematical truth about, say, real numbers, or about polynomials, which is behind every symbol we push. And we need to teach these truths to make algebra meaningful and exciting to the student. Thus, a new program must aim not only at the usual skills but also at an understanding and appreciation of the structure of the real number system, and to a lesser extent, of polynomials. A multitude of exercises is still absolutely necessary for gaining manipulative facility, but these techniques must be tied to the ideas from which they derive their validity.

The writer of new materials and the teacher of these materials must ask the questions: (of himself, not his students) What is the abstract system of which the set of real numbers with addition and multiplication is a model? What are the structural properties of this system? How do these properties motivate and unify the solutions of equations and operations on algebraic expressions and functions? We shall try to provide some answers in succeeding chapters.

The teaching of algebra not only must give the student a glimpse of the structure of the subject but must also treat the language with great care. Statements which record the properties of a mathematical system depend on concise language. difference between "and" and "or", "if" and "only if", "not" and "none", etc., can mean the difference between understanding and misunderstanding. More of these matters in Chapter 2.

Language also involves choice of descriptive words. Unlike the chemist, who uses long compound words to describe his materials, the mathematician often selects common words to describe uncommon concepts. The teacher should beware of dictionary meanings for words such as rational, real, imaginary, complex, group, ring, field, limit, term, factor, domain, range. When these words are used as mathematical terms, they do not have the meanings commonly ascribed to them.

5. A Program for Elementary Algebra. This study is designed to explain what the writers of the SMSG-F (First Course in Algebra) had in mind and what teachers should keep in mind as they teach the F materials.

We shall be concerned with the precise structure of an abstract system called an ordered field, because this system has as a model the real number system. Then we shall dissect this system into subsystems and examine each in search of its relations to the system and to the other subsystems. In this way we may begin to see what underlies elementary algebra.

Chapter 2

LANGUAGE

1. Sets. Much heat has been generated in arguments concerning the role of sets in teaching elementary algebra. Some would consider a course in algebra to be "modern" if it mentions the word "set"; others maintain that sets are an unnecessary confusion.

Let us take the middle road and agree that the study of sets for their own sake probably does not belong in an elementary course. On the other hand, the simple language of sets can greatly enhance the flow and increase the interaction of topics in algebra.

A set is a deceptively simple concept: It is merely a collection of objects. The objects, or elements, in a set have at least one common characteristic -- the characteristic of belonging to the same set. This is not double talk. For example, if the set is "my family", it is significant to say that a person x belongs to my family. The set of integers

$$\{2, 4, 6, 8\}$$

has four elements each of which happens to be an even integer. But these four numbers constitute a set merely because they have been listed together. The point is that often we are concerned with a set rather than with its individual elements. Thus, a line is a set of points, but we may think of the line as one entity, or even as an element of a set of lines. In fact, much of mathematics deals with sets of sets of sets of sets ... .

Just as it is possible to describe a number with various names, such as

$$5 - 2 = 3 = \frac{6}{2} = \sqrt{9} = \ldots$$

so it is possible to describe a set in different ways:

{2,4,6,8} = {4,8,6,2} = A: the set of even positive integers less than 10 = B: the set of positive multiples of 2 which are less than 9.

The "=" sign means "is" in the same sense that

"$\frac{6}{2} = \sqrt{9}$" means "$\frac{6}{2}$ is the same number as $\sqrt{9}$";

and for sets A and B,

"A = B" means "A is the same set of elements as B"

If a set A is described by listing its elements in a roster we enclose its elements within braces. If A is described by stating its characteristics we must be certain that the description allows us to determine without ambiguity whether or not an element belongs to the set. "All the whole numbers I can write" does not suffice to define a set unless it is known how much energy and time I have, how long my writing equipment will hold out, and in what order

I propose to write them down. "All the whole numbers greater than 3 and less than 4" does define a set, namely the null or empty set, the set with no elements, symbolized by $\emptyset$.

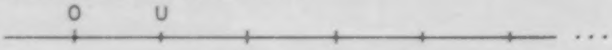
Beginning algebra students bring with them a good deal of information about two sets: the set A of numbers of arithmetic (the non-negative real numbers) and the set P of points on a line. Each of these has interesting subsets.

> If every element of a set S belongs to a set T, then S is a subset of T, and we say that S is contained in T, written $S \subset T$.

Thus, the set W of whole numbers $\{0,1,2,3,\ldots\}$ is a subset of the set A of the numbers of arithmetic: $W \subset A$. It should be noted that a set is always a subset of itself.

> S is a proper subset of T if $S \subset T$, $S \neq \emptyset$ and $S \neq T$.

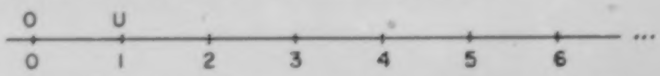
Let us consider a line and two distinct points O and U on the line with U to the right of O. Then take the distance between O and U as a unit measure and mark all points on the line to the right of U and unit distances from each other.



The set M of points so marked is a subset of the set P of all points on the line. An important fact in algebra is that there exists a relationship between the set M of equispaced points and the set W of whole numbers. We say that there is a correspondence between these two sets.

> Given two sets S and T, whenever there is a well-defined rule which associates pairs of elements, the first element of the pair from S and the second element from T, there is a *correspondence* between S and T.

We can define, in fact, a correspondence between M and W which is *one-to-one*; that is, to each point of M we can associate exactly one number of W, and to each number of W exactly one point of M. Let us make the association as in the following figure:



Then we say that each marked point has a corresponding *coordinate*, the whole number associated with the point.

Although it is convenient to speak of these points and numbers interchangeably, such as "the point 2" when we mean "the point whose coordinate is 2", it must be remembered that the set M is *not equal* to the set W. They are quite different sets. But the fact that their elements can be paired off, one-to-one, enables us to carry over to either set the properties of the other.

> A correspondence between S and T is said to be *one-to-one* if each element of S is associated with exactly one element of T and each element of T with exactly one element of S.

We have avoided the necessity of listing a roster of the elements of W by writing

$$W = \{0,1,2,3,\ldots\}$$

to indicate that each element has a successor and, hence, there is no "last element" of W. This is an example of an infinite set, where we intuitively think of "infinity" in terms of "no end".

But intuition is not to be trusted. It would be better to describe an infinite set in some manner which does not involve its elusive "no end". This we shall do as follows. Note that there is a proper subset of W, the set E of all even whole numbers

$$E = \{0, 2, 4, 6, \ldots\},$$

which is in one-to-one correspondence with the set W. We indicate the pairings of elements of this one-to-one correspondence as follows:

$$\begin{array}{ccc} \underline{E} & & \underline{W} \\ 0 & \longleftrightarrow & 0 \\ 2 & \longleftrightarrow & 1 \\ 4 & \longleftrightarrow & 2 \\ 6 & \longleftrightarrow & 3 \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \end{array}$$

This suggests a more satisfactory definition.

> A set T is infinite if there is a proper subset S of T such that S and T are in one-to-one correspondence.

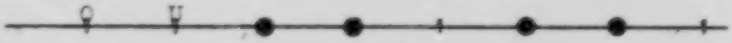
We define a set to be finite if it is not infinite; that is, a finite set cannot be put in one-to-one correspondence with a proper subset of itself.

Now that we have established that the set W of whole numbers is infinite, we can show that the set of points M is infinite as a consequence of the one-to-one correspondence between W and M.

Some infinite sets have the property of being countable, that is, of being in one-to-one correspondence with the set of counting numbers. Later we shall deal with infinite sets which are not countable, such as the set R of all real numbers.

Much of our attention in later chapters will be directed to the fundamental one-to-one correspondence between the set of all the points of a line and the set R of all real numbers. This correspondence is at the heart of coordinate geometry -- the properties of points on the line suggest analogous properties of real numbers, and vice versa.

In order to cash in on this intimate relation between sets of points and sets of numbers we use special language and symbolism to connect them. The number corresponding to a point we have called the coordinate of the point. The set of points corresponding to a certain set of numbers we call the graph of the set of numbers. For example, the set of points indicated by heavy dots in the following figure

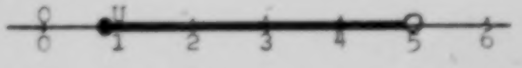


is the graph of the set {2,3,5,6} . The graph of A, where

A = set of all real numbers less than 5 and

greater than or equal to 2/3

s indicated by a heavy line and solid dot in the figure below.



Exercises

1. Given the sets

   $A = \left\{\frac{8}{6}, \sqrt{\frac{9}{4}}, \frac{5}{3}\right\}$

   B = set of all negative integers greater than 3.

   C = set of all whole numbers which are not multiples of 3.

   D = set of all rational numbers between 1 and 2 written with denominators between 1 and 4.

   $E = \{0\}$

   F = set of all rational numbers between 1 and 2 written with numerators between 1 and 6.

   G = set of all whole numbers which are multiples of 3.

   H = set of all numbers x such that $x^2 + x = x$.

   J = set of all numbers of the form $3x + 1$ or $3x + 2$, where x is any whole number.

   Decide which of these sets are equal; which are in one-to-one correspondence.

2. Which of the sets in problem 1 are proper subsets of the set W of whole numbers? of the set G in problem 1?

3. There are various types of correspondences between sets other

than one-to-one. A many-to-one correspondence between S and T associates each element of S with exactly one element of T but at least one element of T with more than one element of S. A many-to-many correspondence is defined accordingly. For each of the following pairs of sets a rule of correspondence is given; decide what type of correspondence it is.

(a) $S = I$ (integers), $T = I$; to each x in S there corresponds $-x$ in T.

(b) $S = I$, $T = I$; to each x in S there corresponds x^2 in T.

(c) $S = R$ (real numbers), $T = R$; to each x in S there corresponds y in T such that $x + y = 7$.

(d) $S = R$, $T = R$; to each x in S there corresponds y in T such that $x^2 = y^2$.

(e) $S = I$, $T = I$; to each x in S there corresponds y in T such that $x^2 = y^3$.

(f) $S = \{1,2,3,\ldots,31\}$, $T = \{\text{Sun., Mon., Tues., \ldots, Sat.}\}$, with the correspondence given by the calendar for July of this year.

4. Which of the following sets are infinite?

(a) Set C of Problem 1.

(b) Set D of Problem 1.

(c) Set of all positive rational numbers written with denominator 3.

(d) Set of all numbers of the form $a\sqrt{2}$, where a is an integer.

5. We may show that a set S is closed under a binary operation $*$ as follows: Construct the set T of all elements of the form $x*y$, where x and y belong to S. If T is a subset of S, then S is closed under $*$. (It is not necessary that T be a proper subset of S.) Decide whether the following sets are

closed under the indicated operations:

(a)	All whole numbers which are not multiples of 3.	multiplication
(b)	All whole numbers which are not multiples of 4.	multiplication
(c)	{0,1}	multiplication
(d)	{0,1}	addition
(e)	All positive integers.	subtraction
(f)	All positive rational numbers.	division
(g)	All positive integers.	half the sum
(h)	All even integers.	half the product
(i)	All squares of integers.	addition
(j)	All rational numbers between 0 and 1.	multiplication

With what mathematical facts do you associate the answers to (a) and (b)?

A unary operation is performed on a single element. Decide whether the sets are closed under the indicated unary operations:

(k)	All positive rational numbers.	square root
(l)	All integers.	squaring
(m)	All even integers.	half the square

6. Draw the graphs of the sets:
   (a) D of Problem 1.
   (b) F of Problem 1.

2. *Sentences*. The properties of an abstract system could be described and recorded in terms of English sentences. But there is much efficiency and avoidance of ambiguity to be gained by abbreviating English sentences into mathematical sentences. Thus we abbreviate the sentence

"Five plus three is nine."

to

$$5 + 3 = 9,$$

meaning, of course, that "$5 + 3$" and "9" are different symbols for the same number. There is no doubt that we have written an English sentence, but there may be some doubt about the corresponding mathematical sentence. It *is* a sentence, even though the statement it makes is false. We shall be concerned with sentences or statements which we assume are either *true* or *false*, but not both, and have meaning and content. Any statement to which this assumption does not apply shall be excluded from our discussion. For example, "$4 =$ a triangle," is without meaning and will not be considered as a sentence. Also, " $3^{+} - (\ ) = 2\sqrt{\ }$ " makes no sense because it does not conform to accepted mathematical grammar. On the other hand, "Every positive even integer is the sum of two primes," is a sentence because even though no one knows whether it is true, we are willing to accept it as either true or false. The assumptions that a sentence is either true or false, but not both, are often called the laws of *contradiction* and the *excluded middle* of logic.

Simple sentences concerning numbers may involve any of the verb symbols $=$, $<$, $>$, $\neq$, $\nless$, $\ngtr$, which have the usual meanings of equality and order and their negations. Compound sentences are constructed from simple sentences by conjunction, disjunction, or conditional.

> If A, B are sentences, then the sentence
>
> A and B (conjunction)
>
> is true if both A and B are true; otherwise it is false. The sentence
>
> A or B (disjunction)
>
> is false if both A and B are false; otherwise it is true.

For example, the disjunction

$$5 < 6 \text{ or } 5 = 6 \quad (\text{abbreviated } 5 \leq 6)$$

is true because at least one of the sentences, namely "$5 < 6$", is true. But the conjunction

$$5 < 6 \text{ and } 5 = 6$$

is false because at least one of the sentences, namely "$5 = 6$", is false.

> If A, B are sentences, then the sentence
>
> if A, then B (conditional)
>
> is false if A is true and B is false; otherwise, it is true.

For example, the conditional

$$\text{if } 2 + 3 = 5, \text{ then } 3 + 4 = 6$$

is false because the sentence A: $2 + 3 = 5$ is true and the sentence B: $3 + 4 = 6$ is false. On the other hand, the conditional

if $2 + 3 = 4$, then $3 + 4 = 7$

is true because "$2 + 3 = 4$" is false and "$3 + 4 = 7$" is true.

At this point it is instructive to list the possibilities which, according to the definition, yield a true conditional.

A	B	if A, then B
True	True	True
False	True	True
False	False	True

The remaining possibility, namely A true and B false, is the only one for which "if A, then B" is false.

At first thought this definition of a conditional seems to violate the common meaning of "if A, then B". Actually, this definition is motivated by our desire to express any valid reasoning leading from a sentence A to a sentence B. Certainly, if A is true, then any reasoning process that is valid will lead us from A to a true conclusion B. This is the first possibility listed in the above table. But we must also acknowledge that if we argue from a false premise A and proceed by means of valid reasoning to a conclusion B, then B may sometimes be true, sometimes false. The emphasis is on the validity of the reasoning. For example, if we take as our premise A: $5 = 4$, we may add 3 to both members to obtain B: $8 = 7$, which is false; we may, instead, remark that "$5 = 4$" and "$4 = 5$" yield B: $5 + 4 = 4 + 5$, which is true. In each case, the reasoning was valid. Hence, it is suggested that our definition of a true conditional include the second and third possibilities in the table. Of course, after we agree on a definition, we must forget the motivation which suggested it and accept the form of the conditional even when there is no apparent relation between the sentences A and B in the sentence "if A, then B".

We do not allow the fourth possibility to occur in a valid reasoning process. Thus we call the conditional false if a true A leads to a false B. This can be summed up by saying that the conditional "if A, then B" is true if A is false *or* B is true; it is false if A is true *and* B is false.

We write the sentence

A, if and only if B (*biconditional*)

as an abbreviation for the conjunction

(if A, then B) and (if B, then A).

Thus, a biconditional is true if both A and B are true or if both A and B are false.

The question arises: Is the following a sentence?

$$x + 3 = 5.$$

The answer depends on the meaning of the symbol x. If we require that x be a symbol for a number without our stating that number specifically, then "$x + 3 = 5$" is an *open* sentence in the sense that the question of its truth is left open until we specify what number x is. The particular set of numbers from which x is to be chosen is called the *domain* of x.

Here we have the first example of a variable; a more detailed discussion will be given in Chapter 6.

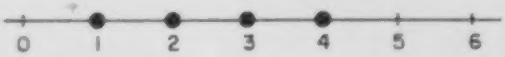
There is a close tie between open sentences in one variable, sets of real numbers, and sets of points on the number line. For example, if the domain of x is the set of all integers, then the open sentence

$$x \geq 1 \text{ and } x < 5$$

(which is usually abbreviated to "$1 \leq x < 5$") is true when x is chosen as any element of the set

$$\{1,2,3,4\}.$$

And this set has the graph

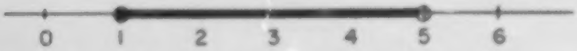


It is natural to call $\{1,2,3,4\}$ the truth set of the sentence and the graph of this set the graph of the sentence.

> The truth set of a sentence in one variable is the set of all numbers in the domain of the variable, and only those numbers, which make the sentence true.

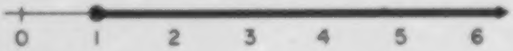
Thus an open sentence in one variable is a sorter which separates the domain of the variable into two subsets, one the truth set of the sentence, and the other the set of the remaining numbers.

Note the importance of specifying the domain of the variable. If for the sentence "$x \geq 1$ and $x < 5$" the domain is, instead, the set of all real numbers, then its graph is

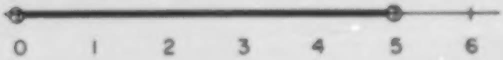


It is instructive to compare the graphs of the three sentences "$x \geq 1$", "$x < 5$", "$x \geq 1$ and $x < 5$", where the domain of x is, say, the set of all positive real numbers.

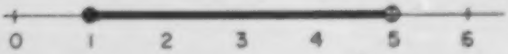
$x \geq 1$



$x < 5$



$x \geq 1$ and $x < 5$

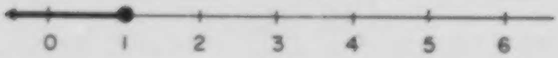


We see that the graph of "$x \geq 1$ and $x < 5$" consists of all the points which are in both the graph of "$x \geq 1$" and the graph of "$x < 5$".

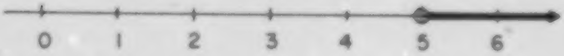
> If S and T are sets, the set of elements each of which belongs to both S and T is the intersection of S and T.

Consider the sentences "$x \leq 1$", "$x > 5$", "$x \leq 1$ or $x > 5$", where the domain of x is the set of all real numbers.

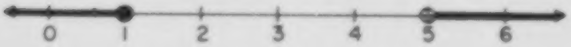
$x \leq 1$



$x > 5$



$x \leq 1$ or $x > 5$



The graph of "$x \leq 1$ or $x > 5$" contains all the points which belong to either the graph of "$x \leq 1$" or to the graph of "$x > 5$".

> The set of elements each of which belongs to either S or T is the union of S and T.

As another example consider the open sentence

if $y \leq 3$, then $y > 5$, y any integer.

The truth set of this open sentence must contain all the integers greater than 3 (since for these integers the sentence "$y \leq 3$" is false); it must also contain all the integers greater than 5 (since for these integers the sentence "$y > 5$" is true). Hence, the truth set is the set of all integers greater than 3.

Consider the open sentences

(1) if $r < 3$, then $4 = 2$, r any integer

and

(2) if $3 = 5$, then $q = 1$, q any integer.

Since in sentence (1), B: $4 = 2$ is false for all integers, the conditional is true only for those integers for which A: $r < 3$ is false, i.e., for $r \geq 4$. In sentence (2), A: $3 = 5$ is false for all integers; hence, the conditional is true for all integers.

A sentence in two ordered variables has a truth set consisting of a set of ordered pairs of numbers, the first number of each pair corresponding to the first variable and the second number to the second variable, such that these pairs and only these pairs make the sentence true. For example, if x (the first variable) and y have as domains the set of positive integers, then the sentence

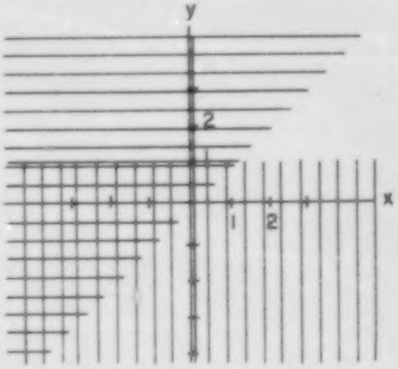
$$x + y = 5$$

has the truth set $\{(1,4), (2,3), (3,2), (4,1)\}$.

The graph of a set of ordered pairs of numbers is the set of points on a plane located with respect to two perpendicular number lines as follows: If the number lines coincide at their 0 points, the number pair (a,b) corresponds to a point P whose projection on the first line has coordinate a and whose projection on the other line has coordinate b. For example, the graph of the sentence

$$x < y \quad \text{and} \quad y < 1,$$

where the domains of x and y are the set of all real numbers, is obtained as follows. The truth set of "$x < y$" is the set of all ordered pairs of real numbers for which the first number is less than the second; the truth set of "$y < 1$" is the set of all ordered pairs for which the second number is less than 1. In the adjacent figure the graphs of the separate sentences are shown with different shadings, and the graph of the conjunction "$x < y$ and $y < 1$" is shown with double shading. (Of course,



the shadings terminate at the edges of the figure because of limitations of space and not because the graphs terminate there.)

These are the building blocks of algebra:

(1) Solving an equation in one variable is nothing more than determining the truth set of the open sentence. The solution* depends on the set of numbers available for the domain of the variable -- for certain domains the solution may be the null set, whereas for other domains its truth set may be non-empty.

(2) Stating a property of an algebraic system is a matter of writing an open sentence which is true for every element of the system. For example, the distributive property of the system of real numbers can be stated as:

> For any real number x and any real number y and any real number z, the sentence
> $$x(y + z) = xy + xz$$
> is true.

This we usually abbreviate to:

> For any real $x, y, z,$
> $$x(y + z) = xy + xz.$$

As another example, let A be a variable whose domain is the set of all sentences. Then the law of contradiction of logic can be stated as:

---

*We often say "solution" for "truth set", particularly if the sentence is an equation.

For every A, the sentence

A or not-A

is true.

And the law of the excluded middle is:

For every A, the sentence

A and not-A

is false.

Here we use the notation "not-A" to denote the negative or denial of A, that is, the sentence which is false when A is true, and true when A is false.

Exercises

1. If T denotes "true" and F "false", fill in the following tables with T or F, if possible, where A and B are sentences.

(a)

A	B	A and B	A or B	not-A	if A, then B	not-A or B
T	T					
F	T					
T	F					
F	F					

(b)

A	B	A and not-B	if not-A, then B	if B, then A
T	T			
T	F			
F	T			
F	F			

(c)

A	B	if A, then B	A or B	A and B	A and not-A
		F			
T			F		
				T	
	F	T			

2. Let the domain of t be the set of positive integers. Then find the truth set of each of the sentences:

(a) $8 + t < 12$ or $5 + 1 \neq 6$

(b) $8 + t < 12$ or $5 + 1 = 6$

(c) $8 + t < 12$ and $5 + 1 \neq 6$

(d) if $8 + t < 12$, then $5 + 1 = 6$

(e) if $5 + 1 \neq 6$, then $8 + t < 12$

(f) if $8 + t < 12$, then $5 + 1 \neq 6$

(g) $t + 2 = 4$ or $t + 2 \neq 4$

(h) $t + 2 \leq 4$ and $t + 3 > 4$

(i) ($t + 2 < 4$ or $t + 2 < 5$) and $t + 2 > 3$

(j) $t + 2 < 4$ or ($t + 2 < 5$ and $t + 2 > 3$)

3. Let the domain of t be the set of real numbers. Then draw the graph of each sentence in Problem 2.

4. Find the truth set of each of the following sentences, where x is the first variable, for the indicated domain R of x and y:

(a) $x = y^2$, $R = \{1,2,3,\ldots,36\}$

(b) $x + 2 = y$ and $x + y = 4$, $R = \{1,2,3,\ldots\}$

(c) $x + y = 5$ or $2x + y = 6$, $R = \{1,2,3,\ldots\}$

(d) $x + y = 5$ and $2x + y = 6$, R = set of all real numbers.

(e) $x < 3$ and $y > 1$, $R = \{1,2,3,4,5\}$

(f) $x + 2y > 0$ and $2x + y < 0$, R = set of all integers between -4 and 4.

5. Draw the graphs of the following sentences for the indicated domain R of each of the variables: (Consider x as the first variable.)

(a) $x + y = 3$ and $2x + y = 5$, R = set of all real numbers.

(b) $x + y = 3$ or $2x + y = 5$, R = set of all real numbers.

(c) $1 < x^2 + y^2 \leq 4$, R = set of all real numbers.

(d) $x < 3$ or $y > 1$, R = set of all real numbers.

(e) $x > 2y$ and $x < 1$, R = set of all real numbers.

(f) $x < y$ and $x < -y$, R = set of all real numbers.

---

3. Logic. In Chapter 3 we shall prove the following property of real numbers:

> For any real numbers a and b, $ab > 0$ if and only if $(a > 0$ and $b > 0)$ or $(a < 0$ and $b < 0)$.

How do we know this sentence is true for any two real numbers a and b? Is this a rule laid down arbitrarily by mathematicians? Or did this property arise through experience with numbers by trial and error?

It happens that this property can be proved as a consequence of other more fundamental properties; that is, it is a theorem. But what about the other properties from which it is deduced? Are they also theorems? This line of questioning would eventually lead us back to a certain basic set of properties of real numbers. No property included in this basic set of properties could then be deduced from the other properties in this set. We are then left with a set of properties that cannot be proved.

Even the words and symbols used in the above property give us trouble. What does the symbol "<" mean? When we define its meaning in terms of other words and symbols, we will again be squeezed back to a set of words and symbols no one of which can be defined in terms of the others (unless we are tempted to define these basic words in terms of words already defined -- a circular kind of definition which is mathematically taboo).

Thus we must begin a study of any mathematical system with a set of undefined words and symbols. Although no attempt is made to define these words formally we always have in mind one or more representations of the words. In a study of plane geometry, for example, we begin with undefined words such as "point", "line", "on", "equal", but we may visualize "point" as a spot of ink on a paper, "line" as a streak of ink, etc. In algebra we can begin with the undefined words "number", "sum", "product", "equal", "less than", and symbols representing these words. It is possible to think of many kinds of "numbers" and sums and products of numbers as representations or models of these words, but any logical deduction from these words must be independent of the particular interpretations that might be attached to them.

It should be understood that the set of words left undefined is somewhat arbitrary and is determined partly by convenience (or convention) and by the amount of rigor demanded. A smaller set of words may be possible, or even a different set. Then the others are defined in terms of this set.

Having decided upon a basic set of undefined words, we next agree upon certain properties that we shall assume these words obey. These properties are stated in forms of open sentences, and they impose conditions upon the undefined words. That is, we do not define the words but we assume they satisfy certain conditions. These assumptions we call axioms. They are not "self-evident" or obvious. They are properties which are assumed to be true. The axioms chosen are often suggested by our experience with the model we had in mind for the undefined words. We must, however, regard the axioms as independent of any empirical considerations. In this way we hope, by deduction, to make discoveries without explicit experience and then to use these new facts as a check on our experience, and vice versa.

After we select a set of axioms, that is, a set of properties which we assume are obeyed by our set of undefined terms, we may then prove theorems. These are sentences which can be proved true in accordance with the laws of logical deduction on the basis of the accepted axioms. In this way we build a body of knowledge about a mathematical system. In summary, first assume a set of axioms to be true. Prove that if the axioms are true, then certain theorems are true. Then prove that if these theorems and the axioms are true, then certain other theorems are true, etc. In the process, from time to time we define certain new words and symbols in terms of the basic set of undefined words and symbols, and then other words in terms of these words, etc. At no point in the process may we use any information other than that obtained in a prior theorem, a prior definition, or the axioms.

Thus, all the procedures and rules of algebra can be stated as theorems which can be derived from a small set of axioms. We shall list these axioms in Chapter 3.

At this point an objection might be raised. Why can't we avoid all this bother and simply take all the results in algebra as rules without worrying which must be proved and which can be assumed? In fact, why prove results that seem obvious anyway?

There is danger inherent in accepting a list of rules without proof. How can a given rule be tested with respect to its validity?

We cannot check its truth for every value of the variables in general. And if the rules are not derived from some basic set of rules, we will never have assurance that they are consistent with each other. We say that two statements are inconsistent if they contradict each other, that is, if they lead to a statement of the form "A and not-A". As noted earlier, such a statement is false for every sentence A.

Note that it is easy to see how to prove the inconsistency of a set of statements; the existence of one counter-example, a specific contradiction, is enough. But proving consistency is another matter. Regardless of how hard one searches for counter-examples, the mere fact that none has been found does not guarantee consistency. When the search is called off, the very next example might have yielded a counter-example. The only way to guarantee consistency of a set of statements is to prove that they are logical consequences of a set of consistent statements, namely, a set of consistent axioms.*

Now that we have discussed the need for proving the results of algebra, the question remains: What do we mean by a "proof"? Too often a result is considered proved if it is "believed", or if it is plausible, or if it is known to be true in a few cases. Having faith in a statement is not enough. In the latter category is the so-called "proof" by induction:** "It has been observed that the result has held true in n trials in the past; hence, it will continue to hold true in all cases."

---

*The problem of proving that a given set of axioms is consistent is a fundamental and difficult job not to be tackled here. Indeed, in some cases possibly it cannot be proved.

**Not to be confused with mathematical induction, which is a powerful and valid method of proving special types of statements. (See Section 4.1.)

We need not belabor the fact that this is not a proof. Of course, by induction one may arrive at a conjecture which can then be proved by deductive methods. G. Polya has written a fascinating set of books on this subject, Mathematics and Plausible Reasoning, Princeton University Press, 1954, in which he investigates how one discovers what statements are worth trying to prove and what suggests such statements in the first place.

In algebra, theorems are written as conditional compound sentences of the form

if p, then q,

where p and q are open sentences. Thus we mean by the proof of a theorem the process of showing that a conditional sentence is true for all values of the variables (let us call such a conditional true, for short). The sentence p is called the hypothesis; it is known or assumed to be true. The sentence q is the conclusion; it must be proved true. There are several methods of proof available.

Direct proof. A basic rule of logic is the law of transitivity of conditionals:

> If the conditionals (if A, then B)
> and (if B, then C) are true, then
> the conditional (if A, then C) is
> true.

This law can be stated more compactly by writing the true conditional "if A, then B", read "A implies B", as

$$A \Longrightarrow B.$$

(A true conditional is often called an implication.) Then the law becomes

$$(A \Longrightarrow B \text{ and } B \Longrightarrow C) \Longrightarrow (A \Longrightarrow C),$$

and the transitivity of the implications is more apparent.

A direct proof of the theorem "$p \Longrightarrow q$" is usually effected by collecting known axioms and theorems in the following format:

$$p \Longrightarrow r, \quad r \Longrightarrow s, \quad s \Longrightarrow t, \quad \ldots, \quad u \Longrightarrow q.$$

Then, by transitivity, $p \Longrightarrow q$.

For example, consider the theorem:

If $a = b$, then $a - b = 0$.

Here we must prove "$p \Longrightarrow q$", where p is the sentence "$a = b$" and q is the sentence "$a - b = 0$". Let us assume it has been established previously that

$$a = b \Longrightarrow a + (-b) = b + (-b)$$

and

$$a + (-b) = b + (-b) \Longrightarrow a - b = 0.$$

Then, by transitivity,

$$a = b \Longrightarrow a - b = 0.$$

In passing, we should note the many forms in which the implication is written in mathematics, all of which are equivalent: (Two open sentences are equivalent if their truth sets are equal.)

(if p, then q) is true

$p \Longrightarrow q$

q, if p

p, only if q

q is a necessary condition for p

p is a sufficient condition for q

if not q, then not p.

The last of these is called the contrapositive of the implication. It states that if q is false, then p must also be false.

As was mentioned before, the sentence "p, if and only if q" is really a statement of the conjunction of two conditionals:

(if p, then q) and (if q, then p).

Thus to prove a theorem of the form "p, if and only if q", we must really prove two theorems:

$p \Longrightarrow q$ and $q \Longrightarrow p$.

Indirect proof. This method is often called "proof by contradiction". By a contradiction we mean a sentence of the form "A and not-A". We assumed earlier (the law of contradiction) that such a sentence is always false.

For example, let us prove the theorem:

> If a is an integer and a^2 is divisible by 2, then a is divisible by 2.

The hypothesis is: $a^2 = 2c$, for some integer c. Let us add to this hypothesis the denial of the conclusion: $a = 2d + 1$, for some integer d. Now our new hypothesis is:

$$a^2 = 2c \text{ and } a = 2d + 1, \text{ for some integers } c \text{ and } d.$$

By squaring both members we have

$$a = 2d + 1 \Longrightarrow a^2 = 4d^2 + 4d + 1 = 2(2d^2 + 2d) + 1.$$

Then

$$a^2 = 2c \text{ and } a = 2d + 1 \longrightarrow a^2 = 2c \text{ and } a^2 = 2(2d^2 + 2d) + 1.$$

The latter sentence is a contradiction because a^2 cannot be twice an integer and at the same time one more than twice an integer. Hence, the assumption that a is not divisible by 2 led to a contradiction; therefore, a is divisible by 2. Here we took as part of our hypothesis the assumption that the conclusion q is false. From this hypothesis we derived a sentence of the form "A and not-A", that is, "A is true and A is false" and we thus proved that our assumption about the conclusion was invalid; hence, q is true.

> The method of indirect proof consists of proving "$p \Longrightarrow q$" by proving that "p and not q" implies a contradiction; that is,
>
> $$(p \text{ and not } q) \Longrightarrow r,$$
>
> where r is a contradiction.

There is no general rule which tells us how to arrive at a contradiction. This comes only with experience. Nevertheless, the indirect method often provides an attack on an "obvious" theorem which eludes the direct method. This is particularly true when the theorem is a statement about all the elements of a set; then an indirect proof deals with some elements not in the set.

Since the statement "not-$q \Longrightarrow$ not-p" is equivalent to "$p \Longrightarrow q$", another indirect method of proof consists of proving the contrapositive of the theorem. In the preceding example, another indirect proof would be obtained by proving the contrapositive, namely, "If the integer a is not divisible by 2, then a^2 is not divisible by 2."

The question pertinent to this study is: How much of algebra should be proved in a first course? A considered opinion is that the student should be asked to prove very few theorems, but that he should be exposed to enough proofs in various degrees of completeness to show him methods and necessity of proof and the joys of devising and understanding theorems. The student should become convinced that it is desirable and necessary for the results he is using to be proved rather than accepted as rules, even though he is not mature enough to carry out such a program in detail. He should be made to realize that one could prove the various properties as consequences of basic properties. He must always be told the truth about a result -- that it can be proved and is being accepted temporarily without proof. Occasionally, the outlines of proofs can be carried out until he eventually acquires a feeling for the meaning of proof. By the end of the course the more able students should be ready for a discussion of the axiomatic basis of algebra. But it is not recommended that such a course be started from a formal list of axioms.

On the other hand, the teacher should have a clear idea from the very beginning of precisely what assumptions underlie the algebra that is being taught. Although it is a long, exacting task to develop all the results of algebra from the axioms, he should be familiar enough with this development to understand its framework and its methodology. Portions of Chapters 3, 4 and 5 will be devoted to this development.

2.32

Exercises

1. Two compound open sentences are equivalent if their truth sets are equal, provided their variables have the same domain. Equivalence can sometimes be shown by means of truth tables. For example, consider the open sentences "if A, then B" and "if not-B, then not-A" (called contrapositives), where A and B are open sentences. For any common value of the variables there are certain possible cases of A, B true or false; for each case we determine the truth of the compound sentences, as follows:

A	B	if A, then B	not-B	not-A	if not-B, then not-A
T	T	T	F	F	T
F	T	T	F	T	T
T	F	F	T	F	F
F	F	T	T	T	T

Since the truth tables for the two compound open sentences are the same in every case, we have shown that the sentences are equivalent. In symbolic form,

$$A \Longrightarrow B \Longleftrightarrow \text{not-}B \Longrightarrow \text{not-}A.$$

By means of truth tables, decide which of the following pairs of open sentences are equivalent.

(a) if A, then B; not-A or B

(b) A or not-B; not-A and B

(c) if A, then B; if B, then A (These sentences are converses.)

(d) if A, then B; if not-A, then not-B (These sentences are inverses.)

(e) not-(A and B); not-A or not-B

(f) not-(A or B); not-A and not-B

(g) not-(if A, then B); if A, then not-B

(h) not-(if A, then B); A and not-B

2. Write in symbolic form the facts about certain pairs of equivalent sentences learned in Problem 1. In particular, what is the negative of a conjunction, of a disjunction, of a conditional? Use these results to write the contrapositive of

   (a) if (A or B), then C

   (b) if A, then (B and C)

   (c) if (if A, then B), then (C or D)

   (d) if (A or not-B), then not-(C or D)

3. Find counter-examples to disprove the statements:

   (a) If x is a real number, then $\sqrt{x^2} + 1 = x + 1$.

   (b) If x is a real number, then $3x^2 + 4 = 4 - 2x + 5x^2$.

   (c) If n is a positive integer, then $n^2 - n + 41$ is a prime.

(d) If x is a real number, then $\sqrt{x^2} = x$.

(e) If x is a real number, then $\frac{x^2 - 1}{x - 1} = x + 1$.

(f) If x is a real number, then $\frac{x}{x} = 1$.

(g) If x and y are non-negative real numbers, then

$$x + y > 2\sqrt{xy}.$$

4. Decide which type of proof is best suited for each of the following theorems. Then prove the theorems.

(a) If the integer a is divisible by 2, then a^2 is divisible by 2.

(b) If a is an integer and a^3 is divisible by 2, then a is divisible by 2.

(c) If b is a prime and b is greater than 2, then b is not divisible by 2.

Chapter 3

STRUCTURE OF THE REAL NUMBER SYSTEM

1. *Axioms for an Ordered Field*. Much of elementary algebra is concerned with the system of real numbers. How does the mathematician study a specific model such as the set of real numbers? He forms an abstract system of undefined terms and operations and assumes that this abstract system obeys the properties that the model is known to possess. Then he studies the abstract system, forgetting that it has any connection with the familiar model. In this way he may discover structural properties that he did not notice in the model.

We shall form an abstract system called a *complete ordered field*. Its elements and operations will be left undefined and its axioms are suggested by our knowledge of real numbers. For convenience, let us call the undefined elements "numbers", remembering that this is merely a name.

Our abstract system then consists of a set $\mathcal{F}$ of undefined elements symbolized by a, b, c, ..., 0, 1, ..., with two undefined binary operations (called addition and multiplication, symbolized by + and ·). In all our work the symbol "=" is an abbreviation for "is" and is used to assert the fact that two particular symbols represent the same element of a set.

The basic properties, axioms, which we shall assume for the system $(\mathcal{F}, +, \cdot)$ are listed in three groups. The first consists of the field axioms; any system with two operations which satisfies these axioms is called a field. Then we shall list the order axioms, which endow the elements of $\mathcal{F}$ with relative size. Finally, the completeness axiom will guarantee that there are enough elements in the system so that it will have all the properties of the real number model. The latter axiom and its implications will be dealt with in Chapter 5.

2. Field Axioms. Let us assume that for any elements a and b in $\mathcal{F}$, there is a unique element $a + b$ and a unique element $a \cdot b$ in $\mathcal{F}$ such that the following are true:

F1 For any a and b in $\mathcal{F}$,

$$a + b = b + a \quad \text{and} \quad a \cdot b = b \cdot a.$$

F2 For any a, b and c in $\mathcal{F}$,

$$(a + b) + c = a + (b + c) \quad \text{and} \quad (a \cdot b) \cdot c = a \cdot (b \cdot c).$$

F3 For any a, b and c in $\mathcal{F}$,

$$a \cdot (b + c) = (a \cdot b) + (a \cdot c).$$

between the operations; multiplication is distributive through addition. Let it be agreed that multiplication is performed before addition, except where indicated otherwise; then we may write the distributive property as

$$a(b + c) = ab + ac.$$

Each of the operations is assumed to have an identity (or neutral) element. In F4 the symbol "0", zero, is used as the identity for addition; in F5 the symbol "1", one, is taken as the identity for multiplication. Note the important assumption in F5 that $0 \neq 1$; this "obvious" fact does not follow from the other axioms and must be assumed.

Finally, we assume that each a in $\mathcal{F}$ has an inverse under addition, $-a$, and that each non-zero b in $\mathcal{F}$ has an inverse under multiplication, $\frac{1}{b}$. "$-a$" is read "the opposite of a". We call "$\frac{1}{b}$" the "reciprocal of b".

Before we begin to deduce new properties of the operations from the field axioms, let us make some remarks about the equality relation. Since the statement "$a = b$" means that a and b are symbols for the same element, then it follows immediately that if a, b, c, d are any elements of some set, then

E1 $a = a$

E2 If $a = b$, then $b = a$.

E3 If $a = b$ and $b = c$, then $a = c$.

E4 If $a = b$ and $c = d$, then $a + c = b + d$.

E5 If $a = b$ and $c = d$, then $ac = bd$.

F4 There is an element in $\mathcal{F}$, denoted by "0", such that

$$a + 0 = 0 + a = a$$

for every a in $\mathcal{F}$.

F5 There is an element in $\mathcal{F}$, denoted by "1", different from 0, such that

$$a \cdot 1 = 1 \cdot a = a$$

for every a in $\mathcal{F}$.

F6 For each a in $\mathcal{F}$ there is an element in $\mathcal{F}$ denoted by "$-a$" such that

$$a + (-a) = (-a) + a = 0.$$

F7 For each a in $\mathcal{F}$, except 0, there is an element in $\mathcal{F}$ denoted by "$\frac{1}{a}$" such that

$$a \cdot \frac{1}{a} = \frac{1}{a} \cdot a = 1.$$

We note certain familiar properties among the field axioms. First, $\mathcal{F}$ is closed under "+" (*addition*) and under "$\cdot$" (*multiplication*) because we assume for any pair a,b in $\mathcal{F}$ that there is a unique *sum* $a + b$ in $\mathcal{F}$ and a unique *product* $a \cdot b$ in $\mathcal{F}$. (We shall omit the $\cdot$ whenever no confusion is caused and write "ab".)

In F1 we assume that both binary operations are *commutative*, and in F2 that both are *associative*. Up to this point the two operations have symmetrical properties; in fact, they could be interchanged without any effect. But in F3 we assume a connection

These five statements summarize the various consequences of having different symbols available to represent a single element of a set. For example, if a and b represent the same element, and b and c represent the same element, then, because b can represent only one element, these are the same. This fact, stated in E3, is called the transitive property of equality. In the same way, E4 and E5 state the immediate facts that the sum and product of two elements, being unique, cannot be changed by representing them with different symbols. The point to be emphasized is that a given symbol in a given discussion will stand for one and only one element and when various symbols represent the same element any one of these symbols may be substituted for any other without altering the truth of a sentence.

Some writers would prefer to leave the symbol "=" undefined and impose the conditions E1 to E5 as axioms of equality. The same end result is achieved, although this approach is concerned with language rather than mathematics.

The familiar manipulations of algebra are now consequences of F1 to F7. We shall consider a few such consequences (theorems) Some will be proved and others, marked with *, will be left as exercises for the reader. As an example of the style of proof to be used, we prove the theorem: If a, b, and c are any elements of $\mathcal{F}$, then

$$(a + b) + c = b + (c + a).$$

Proof: We want to show that the number obtained as the sum of $(a + b)$ and c is the same number as obtained as the sum of b and $(c + a)$. We know that

$$(a + b) + c = c + (a + b),$$

by applying F1 to $(a + b)$ and c, since $(a + b)$ is in $\mathscr{F}$; also

$$c + (a + b) = (c + a) + b,$$

by F2. Hence,

$$(a + b) + c = (c + a) + b,$$

by the transitive property of equality, E3. Also

$$(c + a) + b = b + (c + a),$$

by applying F1 to $(c + a)$ and b; hence,

$$(a + b) + c = b + (c + a),$$

by E3. This is the desired result.

In subsequent proofs we shall abbreviate the work as in the following:

$$\begin{aligned} (a + b) + c &= c + (a + b), && \text{F1} \\ &= (c + a) + b, && \text{F2} \\ &= b + (c + a), && \text{F1} \end{aligned}$$

where closure under addition and the transitivity of equality are used without mention.

Theorem 3.1 If a and b are any elements in $\mathscr{F}$ such that

$$a + b = 0, \quad \text{then} \quad b = -a.$$

Proof: By hypothesis, $a + b = 0$.

Then

$$(-a) + (a + b) = (-a) + 0, \qquad \text{E4}$$

and

$$\big((-a) + a\big) + b = -a, \qquad \text{F2 and F4}$$

and

$$0 + b = -a, \qquad \text{F6}$$

and

$$b = -a, \qquad \text{F4}$$

Note that F6 assumes the existence of an additive inverse $(-a)$ of a. Then Theorem 3.1 shows that there is only one additive inverse of a. This proof that the additive inverse is unique is the first one presented to students in SMSG-F.

Theorem 3.2 (Cancellation property for addition.)

If a, b, and c are any elements in $\mathcal{F}$ such that $a + b = a + c$, then $b = c$.

Proof: We know that $-a = -a$ and $a + b = a + c$, by hypothesis. Then

$$(-a) + (a + b) = (-a) + (a + c), \qquad \text{E4}$$

$$[(-a) + a] + b = [(-a) + a] + c, \qquad \text{F2.}$$

Then

$$0 + b = 0 + c, \qquad \text{F6}$$

$$b = c, \qquad \text{F4.}$$

Theorem 3.3 For any a in $\mathcal{F}$,

$$a = -(-a).$$

Proof: $(-a) + a = 0$, F6.

Hence, $a = -(-a)$, Theorem 3.1.

Theorem 3.4 For any a and b in $\mathcal{F}$,

$$(-a) + (-b) = -(a + b).$$

Proof:

$$\begin{aligned}
(a + b) + [(-a) + (-b)] &= [(a + b) + (-a)] + (-b), && \text{F2}\\
&= [(-a) + (a + b)] + (-b), && \text{F1}\\
&= [\big((-a) + a\big) + b] + (-b), && \text{F2}\\
&= (0 + b) + (-b), && \text{F6}\\
&= b + (-b), && \text{F4}\\
&= 0, && \text{F6.}
\end{aligned}$$

Hence, $(-a) + (-b) = -(a + b)$, Theorem 3.1.

Theorem 3.5 The equation $a + x = b$ has the unique solution $(-a) + b$; that is, there is one and only one number x such that $a + x = b$, namely, $x = (-a) + b$.

Proof: First, we verify that if $x = (-a) + b$, then $a + x = b$; that is, we verify that $(-a) + b$ is a solution. If $x = (-a) + b$, then

$$\begin{aligned}
a + x &= a + [(-a) + b], && \text{E4}\\
&= [a + (-a)] + b, && \text{F2}\\
&= 0 + b, && \text{F6}\\
&= b, && \text{F4.}
\end{aligned}$$

Next, we show that this solution is unique. Suppose there are two solutions x and x'. Then

$$a + x = b \quad \text{and} \quad a + x' = b,$$

and $a + x = a + x'$, E3

$x = x'$, Theorem 3.2.

Thus, there is only one solution.

In these proofs we see the interesting and powerful way in which we use the uniqueness of the additive inverse.

Each of the Theorems 3.1 to 3.5 involves only the operation of addition. Corresponding theorems involving multiplication are proved by the simple device of replacing $+$ by $\cdot$, 0 by 1, $-a$ by $\frac{1}{a}$. Notice the parallel among the following: (a and b any elements in $\mathcal{F}$)

Theorem 3.1	Theorem 3.1'
$a + b = 0 \Longrightarrow b = (-a)$	$ab = 1 \Longrightarrow b = \frac{1}{a}$
Theorem 3.2	*Theorem 3.2'
$a + b = a + c \Longrightarrow b = c$	$ab = ac$ and $a \neq 0 \Longrightarrow b = c$
Theorem 3.3	*Theorem 3.3'
$a = -(-a)$	$a = \dfrac{1}{\frac{1}{a}}$, $a \neq 0$
Theorem 3.4	*Theorem 3.4'
$(-a) + (-b) = -(a + b)$	$(\frac{1}{a})(\frac{1}{b}) = \frac{1}{ab}$, $ab \neq 0$

Theorem 3.5

$a + x = b$ has the unique solution $(-a) + b$.

*Theorem 3.5'

$ax = b$, $a \neq 0$, has the unique solution $(\frac{1}{a})b$.

We shall give the proof of Theorem 3.1' and leave the rest for the reader. If $ab = 1$, then $a \neq 0$ (why?) and

	$(\frac{1}{a})(ab) = (\frac{1}{a})\cdot 1$,	E5.
On the left,	$(\frac{1}{a})(ab) = (\frac{1}{a}\cdot a)(b)$,	F2
	$= 1\cdot b$,	F7
	$= b$,	F5.
On the right,	$(\frac{1}{a})(1) = \frac{1}{a}$,	F5.
Hence,	$b = \frac{1}{a}$,	E3

Theorem 3.6 For any a in $\mathcal{F}$,

$$a\cdot 0 = 0\cdot a = 0.$$

Proof:	$a = a$ and $1 + 0 = 1$,	F4
	$a(1 + 0) = a\cdot 1$,	E5
	$= a$,	F5
	$= a + 0$,	F4
Also	$a(1 + 0) = a\cdot 1 + a\cdot 0$,	F3
	$= a + a\cdot 0$,	F5.
Then	$a + a\cdot 0 = a + 0$,	E3
	$a\cdot 0 = 0$,	Theorem 3.1.
Also	$0\cdot a = 0$,	F1.

The reader should not be misled by the apparent duality of the operations and identities which occurs in the two sets of Theorems 3.1 to 3.5 and 3.1' to 3.5'. This duality is not a universal property of the elements of a field, as can be seen by forming a statement corresponding to Theorem 3.6, namely, $a + 1 = 1$, which is certainly not true for some elements in $\mathscr{F}$.

*Theorem 3.7 For any a, b and c in $\mathscr{F}$,

$$(a + b)c = ac + bc.$$

Theorem 3.8 For any a and b in $\mathscr{F}$,

$$(-a)(b) = -(ab).$$

Proof:

$$\begin{aligned} ab + (-a)(b) &= [a + (-a)]b, && \text{Theorem 3.7} \\ &= 0\cdot b, && \text{F6} \\ &= 0, && \text{Theorem 3.6.} \end{aligned}$$

Hence, $(-a)(b) = -(ab)$, Theorem 3.1.

*Theorem 3.9 For any a and b in $\mathscr{F}$,

$$(-a)(-b) = ab.$$

*Theorem 3.10 For any a in $\mathscr{F}$, $(-1)\cdot a = -a$.

New operations on elements of a field are now defined in terms of addition and multiplication.

Definition. For any a and b in $\mathscr{F}$
(1) the number $a - b$, called "the result of subtracting b from a" or "the difference of a and b" is defined as

$$a - b = a + (-b);$$

(2) the number $\frac{a}{b}$, $b \neq 0$, called "the result of dividing a by b" or "the quotient of a and b", is defined as

$$\frac{a}{b} = a \cdot \left(\frac{1}{b}\right).$$

The difference $a - b$ is unique because it is defined as the sum of a and $(-b)$, which is unique. The quotient $\frac{a}{b}$, $b \neq 0$, is unique because it is the product of a and $\frac{1}{b}$, which is also unique.

Here we can quickly clear up the question of division by 0. By Theorem 3.5' the equation $ax = b$ has the unique solution $(\frac{1}{a})b$, which by definition is $\frac{b}{a}$. Now let $b = 1$ and $a = 0$ so that $\frac{b}{a} = \frac{1}{0}$. Assume that $\frac{1}{0}$ is the solution of $0x = 1$. But by Theorem 3.6, $0x = 0$ for every x in $\mathscr{F}$. Hence, the equation $0x = 1$ has no solution. Thus, $\frac{1}{0}$ is not a symbol for any element in $\mathscr{F}$, and a similar argument shows that $\frac{b}{0}$ is not uniquely defined for any b in $\mathscr{F}$. In other words, we have shown that 0 has no reciprocal in $\mathscr{F}$. In our subsequent development we shall always assume that for an element of the form $\frac{a}{b}$, $b \neq 0$.

Among the many theorems that can be proved, we choose as examples the following:

*Theorem 3.11

$$(b - a) + a = b$$

*Theorem 3.11'

$$\left(\frac{b}{a}\right)a = b$$

*Theorem 3.12

$a - (b + c) = (a - b) - c$

*Theorem 3.12'

$\frac{a}{bc} = (\frac{a}{b})(\frac{1}{c})$

Theorem 3.13 $a(b - c) = ab - ac$

Proof: Since $a = a$ and $b - c = b + (-c)$,

$a(b - c) = a\,[b + (-c)]$, E5

$= ab + a(-c)$, F3

$= ab + [-(ac)]$, Theorem 3.8

$= ab - ac.$

*Theorem 3.14 If $ab = 0$, then $a = 0$ or $b = 0$. (Do not confuse with the converse, "If $a = 0$ or $b = 0$, then $ab = 0$," which is a restatement of Theorem 3.6.)

*Theorem 3.15

$a - b = c - d$ if and only if $a + d = b + c$.

*Theorem 3.15'

$\frac{a}{b} = \frac{c}{d}$ if and only if $ad = bc$.

*Theorem 3.16

$(a-b)+(c-d) = (a+c)-(b+d)$

*Theorem 3.16'

$(\frac{a}{b})(\frac{c}{d}) = \frac{ac}{bd}$

Theorem 3.17

$$\frac{ab}{cb} = \frac{a}{c}$$

Proof:

$(ab)c = a(bc)$, F2

$= (bc)a$, F1

$= (cb)a$, F1.

Since $(ab)c = (cb)a$,

$$\frac{ab}{cb} = \frac{a}{c}, \qquad \text{Theorem 3.15'.}$$

Theorem 3.18

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}$$

Proof:

$$\frac{a}{b} = \frac{ad}{bd} \text{ and } \frac{c}{d} = \frac{bc}{bd}, \qquad \text{Theorem 3.17, F1}$$

$$\frac{a}{b} + \frac{c}{d} = \frac{ad}{bd} + \frac{bc}{bd}, \qquad \text{E4}$$

$$= ad\left(\frac{1}{bd}\right) + bc\left(\frac{1}{bd}\right), \quad \text{Definition}$$

$$= (ad + bc)\left(\frac{1}{bd}\right), \quad \text{Theorem 3.7}$$

$$= \frac{ad + bc}{bd}, \quad \text{Definition.}$$

*Theorem 3.19

$$\frac{\frac{a}{b}}{\frac{c}{d}} = \frac{ad}{bc}$$

Exercises

If a, b, c, d are in $\mathcal{F}$, prove the statements of Problems 1 to 7.

1. $(-b) + (a + b) = a$
2. $(\frac{1}{b})(ab) = a$
3. $(-1) + b(a + \frac{1}{b}) = ab$
4. If $a = b$, then $-a = -b$.
5. $-0 = 0$
6. If $a = b$ and $c = d$, then $a - c = b - d$.

7. $a - (-b) = a + b$.

8. Solve the equation: $x + a = b + c$ for x; a, b, c, x in $\mathscr{F}$.

9. Solve the equation: $\frac{x - 1}{x + 1} = 0$ for x in $\mathscr{F}$.

10. Is the binary operation of subtraction commutative? Associative? If not, give counter-examples.

11. A field is an abstract system with two binary operations which satisfies the field axioms F1 to F7. Verify that a commutative group under addition whose non-zero elements form a commutative group under multiplication such that multiplication is distributive through addition, is a field.

12. Consider the set of integers $\{0,1,2,3,4\}$ obtained as remainders after dividing any integers by 5. Define the sum $a + b$ and the product ab of two elements of this set as the remainders after dividing the usual sum and product by 5. Thus, $3 + 4 = 2$, $4 \cdot 4 = 1$, etc. Decide whether this set and these two operations is a field. If so, what is the additive inverse of 3? The multiplicative inverse of 3?

13. Same as Problem 12, except divide by 6.

14. Verify that the set $\{E,O\}$ and the operations $+$ and $\times$ as defined in Problem 1 on page 1.15 forms a field.

15. Prove that in a field the identities for addition and multiplication are unique.

16. Prove that 0 is the only element in $\mathscr{F}$ with the property that $0 = -0$.

17. Consider the system $(S, \oplus, \otimes)$, where S is the set of all elements (a,b), a and b integers, and

$$(a,b) \oplus (c,d) = (a + c,\ b + d)$$
$$(a,b) \otimes (c,d) = (ac, bd)$$
$$(a,b) = (c,d) \Longleftrightarrow a = c \text{ and } b = d.$$

Show that $(0,0)$ and $(1,1)$ are identities for $\oplus$ and $\otimes$, respectively. Is $(S, \oplus, \otimes)$ a field?

---

3. Order Axioms. A second, additional, set of axioms will be assumed for the elements of $\mathcal{F}$. These will impose on the elements of $\mathcal{F}$ an order. A field which satisfies the following axioms will be called an ordered field.

In addition to the relation of equality, we introduce a new relation denoted by the symbol "<". The statement "$a < b$" is read "a is less than b". We assume now that the system $(\mathcal{F}, +, \cdot)$, in addition to being a field, has an order relation "<" and satisfies the

Order axioms

O1 If a and b are any elements in $\mathcal{F}$, then one and only one of the following is true:

$$a = b, \quad a < b, \quad b < a.$$

02 If a, b, c are any elements in $\mathscr{F}$ such that $a < b$ and $b < c$, then $a < c$.

03 If a, b, c are any elements in $\mathscr{F}$ such that $a < b$, then

$$a + c < b + c.$$

04 If a, b, c are any elements in $\mathscr{F}$ such that $a < b$, and $0 < c$, then

$$ac < bc.$$

As a matter of notation we agree that "$a < b$" and "$b > a$" are the same statement, where the latter is read "b is greater than a". Thus the order axioms may be rephrased in terms of the relation "$>$". We also recall that

$$a \leq b \text{ means } a < b \text{ or } a = b$$

and

$$a \nless b \text{ means } a \text{ is not less than } b.$$

In the light of 01,

$$a \nless b \text{ means } a = b \text{ or } a > b$$

and

$$a \nleq b \text{ means } a > b.$$

We say that "a is positive" when $a > 0$ and "a is negative" when $a < 0$.

Note that the truth of "$a < b$" is not altered by using different symbols to represent a and b. This fact can be stated formally as

E6 If $a = c$, $b = d$ and $a < b$, then $c < d$.

Some consequences of the order axioms are stated in the following theorems:

Theorem 3.20 For any a and b in $\mathcal{F}$,

$$a < b \text{ if and only if } 0 < b - a.$$

Proof of $a < b \Longrightarrow 0 < b - a$:

$a < b$,	Hypothesis
$a + (-a) < b + (-a)$,	O3
$0 < b - a$,	Definition, F6.

Proof of $a < b \Longleftarrow 0 < b - a$:

$0 < b - a$,	Hypothesis
$0 + a < (b - a) + a$,	O3
$0 + a < b + [(-a) + a]$,	Definition, F2
$a < b$,	F6, F4.

***Corollaries** to Theorem 3.20

(1) $a > b$ if and only if $a - b > 0$.

(2) $b < 0$ if and only if $-b > 0$.

(3) $b > 0$ if and only if $-b < 0$.

By means of O1 we can sort all real numbers into three disjoint subsets:

(1) The set P of all positive real numbers.

(2) $\{0\}$

(3) The set N of all negative real numbers.

Hence, given any element in $\mathcal{F}$, it belongs to one and only one of the sets P, $\{0\}$, N.

By virtue of O2 to O4 we are assured that if a and b are both positive, then so are $a + b$ and ab. If one is positive and the other is negative, then their product is negative. If both a and b are negative, then $a + b$ is negative and ab is positive. Stated formally, we have

Theorem 3.21 If $a > 0$ and $b > 0$, then

$$a + b > 0 \quad \text{and} \quad ab > 0.$$

Proof:

$a > 0$ and $b = b$,	Hypothesis
$a + b > 0 + b$,	O3
$a + b > b$,	F4
$b > 0$,	Hypothesis
$a + b > 0$,	O2.

Also, if $a > 0$ and $b > 0$, then

$ab > 0\cdot b$,	O4
$ab > 0$,	Theorem 3.6.

***Theorem** 3.22 If $a > 0$ and $b < 0$, then $ab < 0$.

***Theorem** 3.23 If $a < 0$ and $b < 0$, then

$a + b < 0$ and $ab > 0$.

Theorem 3.24 If $a + c < b + c$, then $a < b$.

Proof:

$a + c < b + c$,	Hypothesis
$0 < (b+c) - (a+c)$,	Theorem 3.20
$(b+c) - (a+c) = b - a$,	(why?)
$0 < b - a$,	E6
$a < b$,	Theorem 3.20.

Note that Theorem 3.24 is the converse of O3.

***Theorem** 3.25 If $ac < bc$ and $c > 0$, then $a < b$.

Note the relation between Theorem 3.25 and O4.

***Theorem** 3.26 For $c < 0$,

$a < b$ if and only if $ac > bc$.

***Theorem** 3.27 If $a \neq 0$, then $a^2 > 0$.

It was assumed in F5 that $1 \neq 0$. Now we can use Theorem 3.27 to prove

Theorem 3.28 $1 > 0$

Proof:

$1\cdot 1 = 1$,	F5
$1 \neq 0$,	F5
$1\cdot 1 > 0$,	Theorem 3.27
$1 > 0$,	E6.

Theorem 3.29

$$ab > 0 \Longleftrightarrow (a > 0 \text{ and } b > 0) \text{ or } (a < 0 \text{ and } b < 0)$$

Proof of $\Longrightarrow$: $ab > 0$, by hypothesis. If either a or b is zero, then $ab = 0$, which is contrary to hypothesis. Hence, neither a nor b is zero. If either a or b is positive, say $b > 0$, then if $a < 0$, $ab < 0$, contrary to hypothesis. Hence, if either a or b is positive, the other is also positive. Finally, if either a or b is negative, say $b < 0$, then if $a > 0$, $ab < 0$, contrary to hypothesis. Thus, if either a or b is negative, the other is also negative.

Proof of $\Longleftarrow$: This follow immediately from Theorems 3.21 and 3.23.

*Theorem 3.30

$$ab < 0 \Longleftrightarrow (a > 0 \text{ and } b < 0) \text{ or } (a < 0 \text{ and } b > 0)$$

*Theorem 3.31

$$\frac{a}{b} > 0 \Longleftrightarrow ab > 0$$

The order axioms and the consequent theorems form a basis for the solutions of inequalities. We shall illustrate with several examples.

Example 1. Find the truth set of

$$2x + 4 < 5x + 3, \quad x \text{ in } \mathscr{I}.$$

We know that

$$a < b \Longleftrightarrow a + c < b + c$$

by O3 and Theorem 3.24. Hence

$$2x + 4 < 5x + 3 \Longleftrightarrow -3x < -1$$

by adding $-5x - 4$ to both sides of the inequality. This means that "$2x + 4 < 5x + 3$" and "$-3x < -1$" have the same truth set. We also know from Theorem 3.26 that for $c < 0$,

$$a < b \iff ac > bc.$$

Then

$$-3x < -1 \iff x > \frac{1}{3}$$

by multiplying by $(-\frac{1}{3})$. By transitivity of inequalities, O2, the sentences "$2x + 4 < 5x + 3$" and "$x > \frac{1}{3}$" have exactly the same truth set. Obviously, the solution is the set of all x in $\mathscr{F}$ such that $x > \frac{1}{3}$.

Example 2. Solve $(x - 3)(2 - 3x) > 0$, x in $\mathscr{F}$.

By Theorem 3.29,

$$\begin{aligned}(x-3)(2-3x) > 0 &\iff (x-3 > 0 \text{ and } 2-3x > 0) \text{ or } (x-3 < 0 \text{ and } 2-3x < 0)\\ &\iff (x > 3 \text{ and } x < 2/3) \text{ or } (x < 3 \text{ and } x > 2/3)\end{aligned}$$

The set of elements in $\mathscr{F}$ that are both greater than 3 and less than 2/3 is the null set. Hence, the desired truth set is the set of all x in $\mathscr{F}$ such that

$$x < 3 \text{ and } x > 2/3,$$

that is, such that

$$2/3 < x < 3.$$

Example 3. Solve

$$\frac{x + 1}{1 - x} < 0, \quad x \text{ in } \mathscr{F}.$$

By Theorems 3.30 and 3.31,

$$\frac{x+1}{1-x} < 0 \iff (x+1>0 \text{ and } 1-x<0) \text{ or } (x+1<0 \text{ and } 1-x>0)$$
$$\iff (x > -1 \text{ and } x > 1) \text{ or } (x < -1 \text{ and } x < 1)$$
$$\iff (x > 1) \text{ or } (x < -1)$$

Hence, the solution is the set of all real numbers less than -1 or greater than 1.

By Theorem 3.27 we know that $a^2 \geq 0$ for every a in $\mathscr{F}$. Let us denote the non-negative element whose square is a^2 by the numerals

$$\sqrt{a^2} = |a|.$$

The symbol $\sqrt{a^2}$ is read "the principal (non-negative) square root of a^2" and $|a|$ is read "the absolute value of a". Since $|a| \geq 0$ for every a in $\mathscr{F}$, we may write

$$a \geq 0 \implies |a| = a, \quad a < 0 \implies |a| = -a.$$

> Remark: Two common errors occur frequently on papers of algebra students:
>
> (1) $\sqrt{x^2} = x$, (2) $\sqrt{4} = \pm 2$
>
> The first statement is true only when $x \geq 0$. If $x = -2$, for example, then the statement would read $\sqrt{4} = -2$; but $\sqrt{4}$ is by definition a non-negative number. The student may then argue that $\sqrt{4}$ is either 2 or -2, as in the second statement. The explanation is that every numeral represents exactly one number, and we define $\sqrt{4}$ to mean 2. Then the student may reply that the square of either 2 or -2 is 4. We agree, and we designate the non-negative square root with the symbol $\sqrt{4}$ and the negative square root by $-\sqrt{4}$.

Example 4. Solve $|x - 2| < 1$, x in $\mathscr{R}$.

By definition, $x - 2 \geq 0 \Longrightarrow |x - 2| = x - 2$,

$$x - 2 < 0 \Longrightarrow |x - 2| = -(x - 2).$$

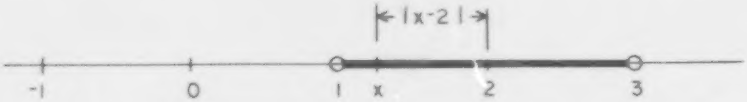
Hence, "$|x - 2| < 1$" is equivalent to

$$x - 2 < 1 \text{ and } -(x - 2) < 1,$$

that is, to

$$x < 3 \text{ and } x > 1, \text{ written } 1 < x < 3.$$

On the number line, $|x - 2|$ represents the distance between the points x and 2. If this distance is to be less than 1, then x must be a point between 1 and 3.



Exercises

1. Use O4 and the definition of "$>$" to verify the statement: If $a > b$ and $0 < c$, then $ac > bc$.
2. Prove that if $a \leq b$ and $b \leq a$, then $a = b$.
3. Prove: $a < b$ if and only if there is a positive number c such that $a + c = b$. This property clearly relates "$<$" to the operation of addition.
4. Prove: For $a > 0$ and $b > 0$,

   $a^2 > b^2$ if and only if $a > b$.

   (Hint: $a^2 - b^2 = (a - b)(a + b)$.)
5. Prove: If $a < b$, then $a < \frac{a + b}{2} < b$.
6. Prove: If $a \neq 2b$, then $\frac{a^2}{4} + b^2 > ab$.
7. Prove: If $a > 0$, then $\frac{1}{a} > 0$.

8. Solve the inequalities: (x in $\mathscr{F}$)

   (a) $2x - 4 < 4x - 7$

   (b) $5x \geq -3x + 8$

   (c) $\frac{2}{x - 1} > 3$

   (d) $(3 - x)(2x - 1) < 0$

   (e) $\frac{3 + x}{2x + 1} > 0$

   (f) $-1 < \frac{x - 1}{x + 1} < 3$

   (g) $x(x + 2)(2 - x) > 0$

9. Find the truth sets of the following sentences, where the domain of x is $\mathscr{F}$.

   (a) $|x - 2| + 1 = 0$

   (b) $x(x - 1) > 0$

   (c) $|x - 3| > 2$

   (d) $|x + 1| + 2 = 3$

   (e) $x^2 - 4x + 3 < 0$

   (f) $|x - 2| > 0$

   (g) $x^4 - 8x^2 + 15 > 0$

   (h) $|x - 1| < 1$ and $|x + 1| > 1$

   (i) $x^2 + 1 \geq 2x$

10. Find the truth set of each of the following sentences, where x is the first variable, for the indicated domain R of x and y.

   (a) $|x| + y = 4$, R = set of all integers greater than -4.

   (b) $|x + y| = 4$, R = set of all integers greater than -4.

   (c) $|x| + |y| = 4$, R = set of all integers.

(d) $|x| = 3$ and $x + y = 1$, $R = \{1,2,3,\dots\}$

(e) $|x| < 3$ and $x + y = 1$, $R = \{1,2,3,\dots\}$

(f) $(x + y)(y - 2) < 0$, $R = \{0,1,2,3\}$

11. Draw the graphs of the following sentences for the indicated domain R of each of the variables. (Consider x as the first variable.)

(a) $|x| + y = 4$, R = set of all real numbers.

(b) $|x| < 3$ and $|y| \leq 2$, R = set of all integers.

(c) $|x| + |y| < 1$, R = set of all real numbers.

(d) $xy > 0$, R = set of all integers.

(e) $(2x - y)y < 0$, R = set of all real numbers.

(f) $|x - y|\,(x - 1) > 0$, R = set of all real numbers.

(g) $(x^2 + y^2 - 4)(x - y) < 0$, R = set of all real numbers.

12. Consider the field described in Problem 12 on page 3.15. If the usual order for positive integers is taken as the order for this field, verify by finding counter-examples that the order axioms are not satisfied.

13. Define an order $\langle$ for $\mathcal{F}$ by the definition

$$a \langle b \Longleftrightarrow |a| < |b|, \quad a, b \text{ in } \mathcal{F}.$$

With this definition, do the order axioms hold?

14. Let us take a different set of order axioms:

There exists a subset of $\mathcal{F}$ called positive numbers such that

(1) for each a in $\mathcal{F}$ exactly one of the following is true:

a is positive, $a = 0$, $-a$ is positive;

(2) if a, b are positive, so are $a + b$ and ab. Now define $a > 0$ to mean "a is positive" and $a < b$ to mean "$b - a$ is positive". With these axioms, prove that O1 to O4 follow as theorems.

---

4. Development of the Real Numbers in SMSG-F. The preceding sections outlined an abstract development of the structure of an ordered field from a set of axioms. The real numbers are a model of such an ordered field; that is, the elements of $\mathcal{F}$ may be identified with real numbers. The rational numbers also satisfy the ordered field axioms. This sort of development gets to the heart of the question: What is elementary algebra made of? It is the kind of development that a teacher of algebra should experience and understand, if possible.

But how should the properties of the real number system be presented to students in a first course? Certainly not abstractly from a set of axioms.

The writers of SMSG-F assumed that a typical student brings to a first course a fairly extensive set of facts about the non-negative real numbers, the so-called numbers of arithmetic. To be sure, his knowledge may include very little about irrationals; he may be able to perform the operation of extracting the square root of 12 to a given number of digits, but here he probably thinks in terms of an operation on an integer rather than the approximation of an irrational number. Considering only the numbers of his experience, he is led to "discover" the commutative, associative and distributive properties. These are then assumed true for all numbers of arithmetic. The same procedure establishes the properties of 0 and 1 and the four order properties. In the process, he experiences a good deal of review and reexamination of arithmetic. The "discovery" is enhanced by associating the non-negative reals with the points of a half-line, and the operations and order relations are given geometric meanings in terms of points on the number line.

Now the stage is set for the crucial step. In order to complete the picture of the real number system, the negative reals must somehow be introduced and their properties established. It is quite natural for a student to accept a set of numbers corresponding to points on the left half of the line. It is also natural to select definitions of operations on these new numbers in such a way that they formally satisfy the same kind of properties enjoyed by the non-negatives.

This presents a problem. Is there a way to define addition and multiplication of the new numbers so that the total set obeys the axioms of an ordered field? Is there only one way?

We first show that there is a way to do this.

The negative numbers are defined by a one-to-one correspondence with the positives, marking the negatives at corresponding distances to the left of the 0 point on the line. Then the resulting numbers are found to have the same order properties as the positives with respect to the number line if "greater than" means "to the right of".

Now the student is told that the total set of numbers, which corresponds to the set of all the points on the number line, is the set of real numbers. These numbers obey the order axioms O1 and O2. (As far as the student is concerned, O3 and O4 have no meanings until addition and multiplication of these numbers are defined.) It remains to define $a + b$ and ab for all real numbers a and b in such a way that all the field axioms are satisfied (again, the student does not think in terms of axioms, but of familiar properties). There is no problem when a and b are positive or zero. Students already know how to add, subtract and multiply such numbers. This suggests the possibility of forming general definitions of sum and product in terms of operations on the non-negative numbers. There is a pitfall here that must be coped with. Unfortunately, the set of numbers of arithmetic is not closed under subtraction. So we must use care in applying subtraction in any definition. At this point a student can determine $a - b$ for numbers of arithmetic only if $a \geq b$.

Here the idea of the opposite of a number in introduced. (This word is selected instead of the usual negative, because it does not have ambiguous meanings; too often a student insists that the negative of a number must be a negative number.) The opposite of a real number is defined to be the number on the opposite side of the 0 point on the line at the same distance from 0. Thus the opposite of -4 is 4, of 5 is -5, etc. The opposite of 0 we take to be 0. There is no confusion between the symbol for the opposite and the symbol indicating a negative number. Note that "-3" represents a negative number; it also represents the opposite of "3". Also, "-(-3)" represents the opposite of the opposite of 3 as well as the opposite of -3, both of which are 3.

Now we define the absolute value of a number to be the larger of the number and its opposite:

absolute value of $a = |a| =$ larger of a and $-a$

Since one of the numbers a and $-a$ must be positive or zero, and a positive is larger than a negative, it turns out that $|a| \geq 0$ for any real number a. In other words, for any real number a, $|a|$ is a number of arithmetic. Now we are ready to define $a + b$ for any real numbers a and b.

The student is led into the definition of the sum of two real numbers by having him consider gains and losses in business transactions. Quickly he catches on to the intent of the definition. Then after many exercises he is invited to discover the definition formally, as follows:

If a and b are positive or zero, $a + b$ is the usual sum of numbers of arithmetic.

If $a < 0$ and $b < 0$, then $a + b = -(|a| + |b|)$; that is, the sum is defined as the opposite of the sum of two numbers of arithmetic, hence, a negative number.

Otherwise, if $|a| \geq |b|$, then

$$a + b = \begin{cases} |a| - |b| & \text{if } a > 0 \text{ and } b < 0 \\ -(|a| - |b|) & \text{if } a < 0 \text{ and } b > 0: \end{cases}$$

if $|b| \geq |a|$, then

$$a + b = \begin{cases} -(|b| - |a|) & \text{if } a > 0 \text{ and } b < 0 \\ |b| - |a| & \text{if } a < 0 \text{ and } b > 0. \end{cases}$$

Thus, we have succeeded in defining $a + b$ in each case as a number of arithmetic or its opposite, where the definition has utilized addition, permissible subtraction, and opposites of numbers of arithmetic only.

Now it is a straightforward matter, although too tedious for the student, to show that

$$a + b = b + a \quad \text{and} \quad (a + b) + c = a + (b + c)$$

for all possible cases of a, b positive, negative or zero. Hence, the operation is commutative and associative. Furthermore, according to this definition, for each real number a there is a unique number $-a$, its opposite, such that

$$a + (-a) = 0;$$

also,

$$a + 0 = a, \text{ for every } a.$$

It remains to define ab for any real numbers a and b. Here we are motivated by the fact that we want the commutative, associative and distributive properties to continue to hold. For example, what should be the meaning of $(3)(-2)$? Of $(-2)(-3)$? We know that

$$\begin{aligned} 0 &= 3\cdot 0 \\ &= 3(2 + (-2)) \\ &= 3(2) + (3)(-2) \quad \text{if the distributive property is to hold,} \\ 0 &= 6 + (3)(-2). \end{aligned}$$

We already have that $-a$ is the only number x such that $a + x = 0$.

Hence, we must take $(3)(-2)$ to be -6 if the distributive property is to hold. Continuing,

$$\begin{aligned} 0 &= (-2)(0) \quad \text{if the property } a\cdot 0 = 0 \text{ is to hold,} \\ &= (-2)(3 + (-3)) \\ &= (-2)(3) + (-2)(-3) \quad \text{if the distributive property is to hold,} \\ 0 &= (-6) + (-2)(-3) \quad \text{if the previous result is to hold.} \end{aligned}$$

Hence, we must define $(-2)(-3)$ to be $-(-6)$, that is, 6.

These examples suggest the definition:

$$ab = \begin{cases} |a||b| & \text{if } a > 0 \text{ and } b > 0, \text{ or } a < 0 \text{ and } b < 0 \\ -(|a||b|) & \text{if } a > 0 \text{ and } b < 0, \text{ or } a < 0 \text{ and } b > 0 \\ 0 & \text{if } a = 0 \text{ or } b = 0. \end{cases}$$

Again it is easy, but detailed, to verify that

$$ab = ba \quad \text{and} \quad (ab)c = a(bc)$$

for all possible cases of a, b negative, zero or positive.

Then, with more effort, it can be shown that

$$a(b + c) = ab + ac$$

for all possible cases. Also, $a \cdot 1 = a$ for all a, and for each $a \neq 0$ there is a unique reciprocal $\frac{1}{a}$ such that $a \cdot \frac{1}{a} = 1$. Thus, the set of numbers consisting of the set of numbers of arithmetic and their opposites satisfies the axioms of an ordered field.

It is recommended that teachers carry out some of the proofs mentioned in the preceding paragraphs. The vast majority of students will be willing to accept the fact that such theorems can be proved, but the more alert ones may want to try the proofs themselves and may need some guidance.

We have shown that there is a way to define order and operations on pairs of negative numbers so that when the negatives are attached to the non-negatives the resulting set of real numbers has the desired field properties. It remains for us to show that given the whole set of non-negatives and negatives, along with the known properties of the non-negatives, there is only one way to define order and the operations of addition and multiplication so that the whole set of numbers satisfies the axioms of an ordered field.

Let us be careful here to lay out exactly what is known and what is to be proved. We are given the numbers of arithmetic (the non-negatives) and the meanings of $+$, $\cdot$, $<$ for any pair of these numbers. We then attach to the non-negatives a set of numbers called negative numbers and write

$$\oplus\ ,\ \odot\ ,\ \prec$$

to indicate any arbitrary meanings of addition, multiplication and order of pairs of numbers. The negative numbers are then defined in terms of the positives: Corresponding to each positive number a there is a negative number -a such that $a \oplus (-a) = 0$. We assume further that $\oplus$, $\odot$, $\prec$ have the conventional meanings of +, $\cdot$, < for pairs of non-negatives and that subtraction of non-negatives is defined as

$$a - b = a \oplus (-b) \quad \text{for} \quad 0 \leq b \leq a.$$

With the assumption that the total set of numbers is an ordered field under $\oplus$, $\odot$, $\prec$, what must be the definitions of these symbols?

From the first it is evident from our assumption that for any two numbers of arithmetic, a and b, ($0 \leq a$ and $0 \leq b$), we have

$$a \oplus b = a + b, \quad a \odot b = ab, \quad a \prec b \Longleftrightarrow a < b.$$

That is, $\oplus$, $\odot$, $\prec$ have the conventional meanings for pairs of non-negatives. In particular, note that $0 \prec a \Longleftrightarrow 0 < a$ and $-(-a) = a$.

Now what is the meaning of $x \prec y$ for any real numbers x and y? Consider the case of x and y negative:

$$\begin{aligned} x \prec y &\Longleftrightarrow x \oplus (-x) \prec y \oplus (-x) , && \text{03} \\ &\Longleftrightarrow 0 \prec y \oplus (-x) , && \text{F6} \\ &\Longleftrightarrow 0 \oplus (-y) \prec [y \oplus (-x)] \oplus (-y), && \text{03} \\ &\Longleftrightarrow (-y) \prec [y \oplus (-y)] \oplus (-x), && \text{F1, F2, F4,} \\ &\Longleftrightarrow (-y) \prec (-x) , && \text{F6, F4} \\ &\Longleftrightarrow (-y) < (-x) , && \text{since } 0 \leq (-x),\ 0 \leq (-y). \end{aligned}$$

But by Theorem 3.26, with $c = -1$, we have

$$(-y) < (-x) \Longleftrightarrow x < y,$$

so that

$$x \prec y \Longleftrightarrow x < y.$$

In particular, for any negative number x, its additive inverse, $-x$, is positive and $0 < (-x)$; then

$$0 < (-x) \Longleftrightarrow x \prec 0 \Longleftrightarrow x < 0.$$

Next, consider the case of x negative and y non-negative:

$$\begin{aligned} x \prec y &\Longleftrightarrow x \oplus (-x) \prec y \oplus (-x) , && \text{03} \\ &\Longleftrightarrow 0 \prec y \oplus (-x) , && \text{F6} \\ &\Longleftrightarrow 0 < y + (-x) , && \text{since } 0 \leq (-x),\ 0 \leq y, \\ &\Longleftrightarrow x < y. \end{aligned}$$

Finally, for the case of x non-negative and y negative:

$$\begin{aligned} x \prec y &\Longleftrightarrow x \oplus (-y) \prec y \oplus (-y) , && \text{03} \\ &\Longleftrightarrow x \oplus (-y) \prec 0 , && \text{F6} \\ &\Longleftrightarrow x + (-y) \prec 0 , && \text{since } 0 \leq x \text{ and } 0 \leq (-y), \\ &\Longleftrightarrow x + (-y) < 0. \end{aligned}$$

But $x + (-y)$ is non-negative. This contradiction shows that $x \preceq y$ cannot be true for x non-negative and y negative. Thus, any arbitrary order $\preceq$ must have exactly the meaning of the conventional $<$ for all real numbers.

Next let us show that for any real numbers x and y, $x \odot y$ must have the same meaning as xy:

Consider the case of *x non-negative and y negative*:

$$0 = y \oplus (-y) \;, \qquad \text{F6}$$

$$x \odot 0 = x \odot [y \oplus (-y)] \;, \qquad \text{E5}$$

$$0 = x \odot y \oplus x \odot (-y) \;, \qquad \text{F3}$$

$$0 = x \odot y \oplus (x)(-y), \text{ since } 0 \leq x \text{ and } 0 \leq (-y).$$

Hence,

$$x \odot y = -[(x)(-y)] \;, \qquad \text{F6}$$

$$= -[|x|\,|y|] \;,$$

$$= xy.$$

The case of *x negative and y non-negative* is handled by observing that $x \odot y = y \odot x$ for all real x and y.

There remains only the case of *x and y negative*:

$$0 = y \oplus (-y) \;, \qquad \text{F6}$$

$$x \odot 0 = x \odot [y \oplus (-y)] \;, \qquad \text{E5}$$

$$(x)(0) = x \odot y \oplus x \odot (-y) \;, \qquad \text{F3}$$

$$0 = x \odot y \oplus -[(-x)(-y)] \;, \text{ by the previous result.}$$

Hence,

$$\begin{aligned} x \odot y &= -[-[(-x)(-y)]], && \text{F6} \\ &= (-x)(-y), \\ &= |x||y|, \\ &= xy. \end{aligned}$$

Thus, for all real x and y, $x \odot y = xy$.

There remains the operation $\oplus$. We shall show that it must have the conventional meaning of +. First consider the case in which x and y are negative:

$$\begin{aligned} (x \oplus y) \oplus [(-x) \oplus (-y)] &= [x \oplus (-x)] \oplus [y \oplus (-y)] && \text{, F1, F2} \\ &= 0 \oplus 0 , && \text{F6} \\ &= 0 , && \text{(why?).} \end{aligned}$$

$$\begin{aligned} \text{Hence,} \qquad x \oplus y &= -[(-x) \oplus (-y)] , && \text{F6} \\ &= -[(-x) + (-y)], \text{ since } 0 \le -x,\ 0 \le -y, \\ &= -(|x| + |y|) . \end{aligned}$$

But this is the definition of $x + y$ for x and y negative.

Hence, $x \oplus y = x + y$, for x and y negative.

For the case of x non-negative and y negative we need to recall that if $0 < b \le a$, then $a \oplus (-b) = a - b$; and we must consider two subcases:

(1) $y < 0 \le x$ and $|y| \le x$:

$$\begin{aligned} x \oplus y &= x \oplus (-|y|), && (\text{if } y < 0, \text{ then } y = -|y|) \\ &= x - |y| , \\ &= x + y . \end{aligned}$$

(2) $y < 0 \leq x$ and $x \leq |y|$:

$$\begin{aligned} x \oplus y &= x \oplus (-|y|) , \\ &= -[(-x) \oplus |y|] , \text{ since } -(a \oplus b) = (-a) \oplus (-b) , \\ &= -[|y| \oplus (-x)] , \qquad \text{F1} \\ &= -(|y| - x) , \\ &= x + y . \end{aligned}$$

Finally, for $x < 0 \leq y$ we use the same arguments as above, knowing that $x \oplus y = y \oplus x$ for all real x and y. Thus for any real numbers x and y it must be true that

$$x \oplus y = x + y .$$

Exercises

1. Show that the definition of $|a|$ given on page 3.30 is equivalent to:

$$|a| = \begin{cases} a & \text{if } a \geq 0 \\ -a & \text{if } a < 0. \end{cases}$$

2. Prove: (a) $|ab| = |a||b|$

(b) $|-a| = |a|$

(c) $-|a| \leq a \leq |a|$

(d) If $b > 0$, then $|a| < b \Leftrightarrow -b < a < b$.

(e) $|a + b| \leq |a| + |b|$

(f) If $|a| = |b|$, then $a^2 = b^2$.

3. Solve: (a) $|x - 2| = |4 - x|$

(b) $|3 - x| = 1$

(c) $|3 - x| < 1$

(d) $|x + 2| = x$

(e) $|x - 2| < 1$ and $|x + 1| < 3$

(f) $|x + 2| < 3$ or $|x - 1| \leq 2$

(g) $|x - 2|^2 + |x - 2| = 2$

(h) $|2x + 1| > 2$ and $|x + 2| < 1$

4. By considering all possible cases, prove the commutative properties of addition and multiplication, assuming they are true for non-negatives and using the definitions of addition and multiplication given on page 3.31 and page 3.32.

5. Show that $a \cdot 1 = a$ for every real number a, assuming it true for non-negatives and using the definition of multiplication on page 3.22.

Chapter 4

SUB-SYSTEMS OF THE REAL NUMBERS

Before completing the list of axioms that describe the abstract system whose model is the set of real numbers (this will be done in Chapter 5) let us show why the list is not already complete. To do this, let us take a different view of the real number model. Instead of looking at real numbers en toto, it will be instructive for us to consider certain sub-systems of the reals and study the properties of these smaller systems. In the process we shall find one proper sub-system of the reals which itself is an ordered field; thus, the axioms for an ordered field do not completely distinguish the system of real numbers from one of its proper sub-systems.

1. The Natural Numbers. If we identify the set $\mathscr{F}$ with the set R of real numbers, then R contains the element 1, by F5. We also know by Theorem 3.28 that $0 < 1$; that is, 1 is a positive real number. Then $1 + 0 < 1 + 1$, by O3. The real number "$1 + 1$" is called "2". (This is strictly a definition, an arbitrary new symbol to abbreviate the symbol "$1 + 1$".) Thus, $1 < 2$, by O2, and $0 < 1 < 2$. In the same way we find that $1 + 1 < 2 + 1$, and, abbreviating "$2 + 1$" to "3" we have $0 < 1 < 2 < 3$. This process is continued by abbreviating

"3 + 1" to "4", "4 + 1" to "5", etc., so that

$$0 < 1 < 2 < 3 < 4 < 5 < \dots .$$

The real number obtained by adding 1 to n is called the successor of n.

> The subset N of R consisting of 1 and every real number which is the successor of a number in N, and no other real numbers, is called the set of natural numbers.

Since 1 is in N, so is its successor, 2; since 2 is in N, so is its successor, 3; etc. Thus

$$N = \{1,2,3,\dots\}.$$

The reader should prove to his own satisfaction that there is no greatest element in N. (Show that the assumption of a greatest element leads to a contradiction.)

Note that the natural numbers are ordered, and each is a positive real number. Since there are real numbers which are not natural numbers, such as -1 (Why?) then $N \neq R$; hence, the set of natural numbers is a proper subset of R. The following is a list of some of the properties of N:

(1) Closure. The set N is closed under addition and multiplication. It is not closed under subtraction or division. The reader should verify that (1 - 2) and $\frac{1}{2}$, for example, are not natural numbers.

(2) *Finite induction*. If a set S of natural numbers contains 1 and if S contains $n + 1$ whenever it contains n, then $S = N$. This property of the natural numbers, called *finite induction*, is a direct consequence of the definition of N and describes conditions under which a set of natural numbers contains *all* the natural numbers.

(3) *Well ordering*. Every natural number is greater than or equal to 1; that is, 1 is the *least element* of N. An ordered set, each non-empty subset of which has a least element, is called a *well ordered set*. Hence, the set of natural numbers is well ordered. This property of the natural numbers follows from the principle of finite induction. The proof is left to the reader.

(4) *Unique factorization*. We define a number p in N to be a *prime* if $p > 1$ and if p cannot be written as the product of two natural numbers between 1 and p. It can be shown that every natural number greater than 1 can be written, *in only one way*, as the product of primes.[*] This fundamental property is called *unique factorization*.

The set N is infinite, since there is a proper subset of N whose elements are in one-to-one correspondence with the elements of N. (Describe such a proper subset of N.) We say that a given infinite set is countable or denumerable if it is in one-to-one correspondence with the set N. Thus, for example, we shall show in Section 3 that the set of rational numbers is countable.

---

[*] See page 23 of *What is Mathematics*, by Courant and Robbins, for a careful proof of this property. Notice how the fact that N is well ordered enters into the proof.

As a consequence of these properties we observe, first, that the set N is of little use in the solutions of equations. Since N lacks closure under subtraction, not even the equation $a + x = b$, where a and b are in N, is guaranteed to have a solution in N. On the other hand, its property of finite induction leads to a technique of proof which can be stated as the:

> Principle of Finite Induction. Let $S(n)$ be an open sentence with one variable n. If
>
> (1) $S(1)$ is true; that is, if the sentence is true when $n = 1$, and
>
> (2) $S(k)$ true $\Longrightarrow S(k + 1)$ true; that is, if the truth of the sentence for n any natural number k implies its truth for $n = k + 1$,
>
> then the truth set of $S(n)$ is the set N, the whole set of natural numbers.

Example 1. Prove:

For any natural number n, $n > 0$.

Proof by finite induction:

Let $S(n)$ be the sentence: $n > 0$. Then

$$S(1):\ 1 > 0,$$

$$S(k):\ k > 0,$$

$$S(k+1):\ k + 1 > 0.$$

By Theorem 3.28 we know that $S(1)$ is true. Hence, 1 is in the truth set of $S(n)$. Next we prove that $S(k) \longrightarrow S(k + 1)$.

$k > 0$, by hypothesis,

$k + 1 > 0 + 1$, O3,

$k + 1 > 1$ and $1 > 0 \longrightarrow k + 1 > 0$, O2.

Thus we have also shown that if k is in the truth set of $S(n)$, then so is $k + 1$. Hence, the truth set of $S(n)$ contains all the natural numbers, and the theorem is proved.

Example 2. Prove:

For any natural number n, $2^n > n$.

Let $S(n)$ be the sentence: $2^n > n$. Then

$$S(1):\ 2^1 > 1$$

$$S(k):\ 2^k > k$$

$$S(k+1):\ 2^{k+1} > k+1.$$

First, we observe that $S(1)$ is true. Next, assume $S(k)$ is true and from this deduce that $S(k+1)$ is true.

$2^k > k$, hypothesis

$2 \cdot 2^k > 2k$, O4

Now, $2 \cdot 2^k = 2^{k+1}$ and $2k \geq k + 1$, (Why?)

Hence, $2^{k+1} > k + 1$, O2

This completes the proof.

The fact that a natural number can be factored into primes in only one way is used constantly in arithmetic computations.

For example, since

$$108 = 2\cdot2\cdot3\cdot3\cdot3 = 2^2\cdot3^3 \quad\text{and}\quad 360 = 2\cdot2\cdot2\cdot3\cdot3\cdot5 = 2^3\cdot3^2\cdot5,$$

we may write the sum

$$\frac{1}{108} + \frac{1}{360} = \frac{1}{2^2\cdot3^3} + \frac{1}{2^3\cdot3^2\cdot5} = \frac{2\cdot5}{2^3\cdot3^3\cdot5} + \frac{3}{2^3\cdot3^3\cdot5}$$

$$= \frac{13}{2^3\cdot3^3\cdot5} = \frac{13}{1080}.$$

Notice how the "least common denominator" of the fractions is obtained in terms of the prime factorizations of the denominators.

In general we say that for a and b in N, a is a factor of b if there is some natural number c such that $ac = b$. Thus, 3 is a factor of 12 because $3\cdot4 = 12$. If a is a factor of b, we say that b is a multiple of a. Thus, 12 is a multiple of 3.

If a is a factor of b, we say that a divides b, sometimes written "$a|b$". It follows that

$$a|b \text{ and } a|c \Longrightarrow a|(b + c)$$

and

$$a|b \text{ or } a|c \Longrightarrow a|bc.$$

These, and other results which the reader can prove, are useful in the factorization of polynomials. (See Problems 9-11.)

Many interesting questions about primes have been answered and some still defy solution. For example, it was shown by Euclid that the set P of primes is infinite. The proof is easily obtained by contradiction.* But it is still not known whether every even natural number can be written as the sum of two primes; yet no one has found an even number which cannot be written as such a sum. $42 = 23 + 19$, $68 = 61 + 7$, etc. Also, certain primes occur in pairs as consecutive odd numbers: 3,5; 5,7; 11,13; 17,19; 29,31; etc. It is not known whether the set of such prime pairs is infinite.

---

*See A Mathematician's Apology, by G.H. Hardy, Cambridge, p. 32.

In summary, $N \subset R$, $N \neq R$, and N has the properties of closure under addition and multiplication, finite induction, well ordering, and unique factorization.

Exercises

1. Prove in any way that each of the following is true for every natural number n:

   (a) $1 + 3 + 5 + 7 + \ldots + (2n - 1) = n^2$.

   (b) $1^2 + 2^2 + 3^2 + 4^2 + \ldots + n^2 = \frac{n(n + 1)(2n + 1)}{6}$.

   (c) $n^2 + 1 > n$.

   (d) 2 is a factor of $n^2 + n$.

   (e) 3 is a factor of $n^3 - n + 3$.

   (f) 4 is a factor of $7^n - 3^n$.
   (Hint: $7^{k+1} - 3^{k+1} = 7^{k+1} - 3 \cdot 7^k + 3 \cdot 7^k - 3^{k+1}$)

   (g) $1^3 + 2^3 + 3^3 + \ldots + n^3 = (1 + 2 + 3 + \ldots + n)^2$

2. Is the set N a group under addition? Under multiplication? If not, what is lacking?

3. If 2 is a factor of a natural number n, we say that n is even; otherwise, n is odd. Prove that for natural numbers a and b:

   (a) If a is even and b is odd, then $a + b$ is odd and ab is even.

   (b) If a is odd and b is odd, then $a + b$ is even and ab is odd.

   (c) If a is even, then a^2 is even.

   (d) If a^2 is even, then a is even.

4. Show that if a is in N and if 3 is a factor of a^2, then 3 is a factor of a.

5. Show that if a is in N and b is a real number not in N, then $a + b$ is not in N.

6. Prove there is no natural number between 0 and 1. (Hint: Prove by finite induction that $S(n)$: $n \geq 1$ is true for all natural numbers.)

7. Prove by finite induction that N is well ordered; that is, that every non-empty subset of N has a least element. (Hint: Let $S(n)$ be the sentence "Any set of natural numbers that contains a number less than or equal to n has a least element.")

8. Prove that if N is well-ordered, then it has the property of finite induction. Hint: Let S be a subset of N that contains 1 and contains $n + 1$ whenever it contains n. Let S' be the subset of N containing all elements of N not in S. Show that S' is empty.

9. Prove the following, for a, b, c in N:
   - (a) $a|b$ and $a|c \Longrightarrow a|(b + c)$
   - (b) $a|b$ or $a|c \Longrightarrow a|bc$
   - (c) $a|b$ and $a|(b + c) \Longrightarrow a|c$
   - (d) $a|b$ and $a\nmid(b + c) \Longrightarrow a\nmid c$ ($a\nmid c$ means that a is not a factor of c.) Hint: Prove the contrapositive.
   - (e) For p a prime, $p|bc \Longrightarrow p|b$ or $p|c$.

10. Use the results of Problem 9 and unique factorization to find two natural numbers b, c, if possible, whose
    - (a) product is 24 and sum is 14.

      (Let $bc = 24$ and $b + c = 14$. Now $24 = 2\cdot 2\cdot 2\cdot 3$. Since $2|24$ and 2 is prime, then $2|b$ or $2|c$. If $2|b$ and $2|(b + c)$, then $2|c$. Hence, $2|b$ and $2|c$. Since $3|24$ and 3 is prime, $3|b$ or $3|c$. But $3\nmid(b + c)$; hence, if $3|b$, then $3\nmid c$, and if $3|c$, then $3\nmid b$. Thus, both b and c contain a factor 2, but only one of b, c contains the factor 3. We conclude that either

      $$b = 2\cdot 3 \text{ and } c = 2\cdot 2 \quad\text{or}\quad b = 2\cdot 2\cdot 3 \text{ and } c = 2.$$

      The second of these possibilities gives $b + c = 14$.)
    - (b) product is 72 and sum is 22.

(c) product is 150 and sum is 25.

(d) product is 84 and sum is 24.

11. Use the techniques of Problem 10 to factor the following polynomials, if possible, into polynomials with coefficients in N.

(a) $x^2 + 8x + 12$

(b) $x^2 + 15x + 56$

(c) $x^2 + 45x + 180$

(d) $x^2 + 32x + 252$

---

2. The Integers. The set N can be enlarged by attaching to it the real number 0 and the additive inverse of each of the elements of N.

> The subset I of the real numbers consisting of all the natural numbers, 0, and the additive inverses of all the natural numbers, and no others, is called the set of integers.

The system of integers is ordered; for we know that $2 > 1$ and $-1 < 0$, so that $(-1)(2) < (-1)(1)$ by Theorem 3.26, and, hence, $-2 < -1 < 0$. In the same way we find that $-3 < -2 < -1 < 0$, and in general

$$\ldots < -4 < -3 < -2 < -1 < 0 < 1 < 2 < 3 < 4 < \ldots .$$

But I is not well ordered, since there is a non-empty subset of I which has no least element. Such a subset is the set of negative integers. To show this we need only note that if n is in this subset, then $n - 1$ is also, and $n - 1 < n$.

There are real numbers which are not in I, such as $\frac{1}{2}$. This seems obvious, but it must be proved. The trick is to show that $\frac{1}{2}$ is greater than every negative integer, is not zero, and is less than every positive integer, and, hence, is not in I. To see this, note that $2 > 0 \Longrightarrow \frac{1}{2} > 0$; $1 < 2 \Longrightarrow (1)(\frac{1}{2}) < (2)(\frac{1}{2}) \Longrightarrow \frac{1}{2} < 1$; and $\frac{1}{2} \neq 0$. Hence, $\frac{1}{2}$ is not a negative integer, since it is greater than 0; $\frac{1}{2}$ is not zero; and $\frac{1}{2}$ is not a natural number (positive integer) since it is less than 1. Thus, we conclude that $\frac{1}{2}$ is not in I and that I is a *proper* subset of R: $N \subset I \subset R$, $I \neq R$, $I \neq N$.

The set I with the operation of addition is a system with the properties of a commutative group. The reader should verify this fact after reviewing the definition of a group. The set I with the operations of addition and multiplication moreover is a system which has the properties of a commutative ring. Again the reader should verify this. The ring of integers has the following properties:

(1) *Closure*. The set I is *closed* under addition, subtraction and multiplication. It is not closed under division. Thus, in the transition from N to I we gained the property of closure under subtraction. Now for any a and b in I, there is a unique solution of $a + x = b$ in I, but not of $ax = b$.

(2) *Unique factorization*. Each integer other than $-1, 0, 1$ can be written as the product of primes and 1 or -1 in only one way; that is, integers have the property of *unique factorization*.

(3) *Countability*. The set of integers is *countable*. To show this, we establish a one-to-one correspondence with N in the following manner:

$$\begin{array}{ccc} \underline{I} & & \underline{N} \\ 0 & \longleftrightarrow & 1 \\ 1 & \longleftrightarrow & 2 \\ -1 & \longleftrightarrow & 3 \\ 2 & \longleftrightarrow & 4 \\ -2 & \longleftrightarrow & 5, \text{ etc.} \end{array}$$

(4) *Division algorithm*. If b is any integer, the integers

$$\ldots, -2b, -b, 0, b, 2b, \ldots$$

are *multiples* of b. Given any integer a, it is either equal to one of the multiples of b or it lies between two successive multiples of b. In the latter case, we mean there is an integer c such that (for b positive)

$$bc < a < b(c + 1);$$

that is,

$$a - bc > 0 \quad \text{and} \quad a - bc < b.$$

Thus, we may set $a = bc + r$, where r is an integer such that $0 < r < b$. If a is a multiple of b, then $a = bc + r$, with $r = 0$. If b is any integer, we see that there is an integer c such that

$$a = bc + r, \quad \text{where} \quad 0 \leq r < |b|.$$

Here we have the important division algorithm which guarantees that for any two integers a and b there is an integer c, called the quotient, and a non-negative integer r, called the remainder, such that $a = bc + r$ and $0 \leq r < |b|$. For example, if $a = 21$, $b = -5$, we may write $21 = (-5)(-4) + 1$. If $a = -2$, $b = 5$, then $-2 = (5)(-1) + 3$. If $a = -21$, $b = 7$, then $-21 = (7)(-3) + 0$.

(5) Decimal representation. The division algorithm for integers allows us to write any integer in a decimal representation:* Given a positive integer d, there are integers $c_0, c_1, c_2, \ldots, c_n$ from the set $\{0,1,2,3,\ldots,9\}$ such that

$$d = c_0 + c_1 10 + c_2 10^2 + \ldots + c_n 10^n.$$

The c's are called digits, and n is some natural number or 0. To show that such a representation is always possible, apply the division algorithm to d and 10, obtaining

$$d = d_1 10 + c_0, \quad 0 \leq c_0 < 10.$$

If $d_1 > 9$, then apply the algorithm again to d_1 and 10, giving

$$d_1 = d_2 10 + c_1, \quad 0 \leq c_1 < 10.$$

Continue this process until a quotient d_n less than 10 is obtained (why will this happen in a finite number of steps?):

---

*A representation of a number is a manner of naming the number.

$$d_{n-1} = d_n 10 + c_{n-1}, \quad 0 \leq c_{n-1} < 10,$$

$$d_n = c_n, \quad 0 \leq c_n < 10.$$

Then, upon eliminating $d_1, d_2, \ldots, d_n$ from these equations we have

$$d = c_0 + c_1 10 + c_2 10^2 + \ldots + c_n 10^n.$$

If d is a negative integer, multiply the decimal representation of $|d|$ by -1.

Although it is customary to restrict the digits to integers from the set $\{0,1,2,\ldots,9\}$ we may let the digits be taken from any set of the form $\{0,1,2,3,\ldots,(p-1)\}$, with $p > 1$, and represent d as

$$d = c_0 + c_1 p + c_2 p^2 + \ldots + c_n p^n.$$

For example, in the ternary representation of d we select the digits from $\{0,1,2\}$. Then $11 = 3 \cdot 3 + 2$ and $3 = 3 \cdot 1 + 0$; hence,

$$11 = 2 + 0 \cdot 3 + 1 \cdot 3^2,$$

which we abbreviate to $(102)_{\text{three}}$, read "one-oh-two, base three".

In summary, when we extend the set N to the set I we lose certain properties and gain others. The properties of finite induction and well ordering are lost. On the other hand, we gain the important property of closure under subtraction. It should be noted in passing that computations with, and representations of, natural numbers are extremely cumbersome without the services of the integer 0. Without 0 any form of representation must be accumulative, such as the Roman numerals. Only with the introduction of 0 can representation be positional, that is, in terms of coefficients of powers of some integer. The fact that every integer can be written as a terminating decimal is of great importance in calculations. (By "terminating" we mean only a finite number of coefficient digits is required in the representation.) The study of integers and their properties, called the theory of numbers, is one of the oldest and most fascinating areas of mathematics.

Exercises

1. If for integers we take similar definitions of factor and multiple as for natural numbers, then the integer a is even if there is an integer c such that $a = 2c$.
   (a) Is 0 an even integer?
   (b) If a is an odd integer and b is an odd integer, is ab an odd integer?
   (c) If a is an integer and a^2 is even, prove that a is even.
2. For integers a and b, recall that $a|b$ means "a is a factor of b". Define the greatest common divisor of a and b, written (a,b), as the greatest positive integer d such that $d|a$ and $d|b$. (Note that for any e such that $e|a$ and $e|b$, we have $e|(a,b)$.)

(a) Compute $(-360,90)$; $(30,$[illegible]$4)$; $(73,-162)$.

(b) If $a > 0$, prove that $(ab,ac) = a\cdot(b,c)$.

3. Is the set of negative integers closed under addition? Under subtraction? Under multiplication?

4. By $(326)_{ten}$ we mean $6 + 2\cdot 10 + 3\cdot 10^2$.

   (a) Convert $(4152)_{six}$ to the "ten" scale, that is, to decimal representation.

   (b) Convert $(101100)_{two}$ to decimal representation.

   (c) Convert $(326)_{ten}$ to a representation in the "three" scale.

   (d) Convert $(326)_{ten}$ to the "nine" scale.

5. Is the set I a group under addition? Under multiplication? If not, explain what is lacking.

6. Is the set I with addition and multiplication a field? If not, what is lacking?

7. Explain the usual algorithms for adding and multiplying integers, "carrying", and subtracting integers in terms of decimal representation.

8. With the ordering given by the axioms O1 to O4 the set of integers is not well ordered. Define a different ordering of I for which *I* is well ordered.

9. Consider the set of integers $T = \{4,7,10,13,16,\ldots,3k+1,\ldots\}$. Let us define a prime in this set to be an element that cannot be obtained as the product of two elements in T. Thus,

4, 7, 10, 13, 19, 22, 25, 31, etc., are primes in T, whereas 16, 28, 40, etc., are composites in T. Can every composite in T be factored uniquely into products of primes in T?

---

3. The Rational Numbers. The system of integers can be extended to a larger system as follows. Consider any integer q in I such that $q \neq 0$. Then by F7 there is a number $\frac{1}{q}$ in R. If p is an integer, the product $p(\frac{1}{q}) = \frac{p}{q}$ is a real number, since the set R is closed under multiplication.

> The subset of R consisting of all real numbers that can be represented in the form $\frac{p}{q}$, where p and q are integers, $q \neq 0$, is called the set F of rational numbers.

The adjective "rational" here implies "ratio", and not the usual dictionary meaning "reasonable" or "sensible".

Here we should comment on the uses of the words "fraction" and "rational number". It must be remembered that a fraction is not a number; it is a symbol which represents a number. By definition, a fraction has the form $\frac{x}{y}$, $y \neq 0$, for any x and y; that is, it indicates a division process. In particular, if x and y represent integers, then the resulting fraction represents a rational number. For example, the fraction "$\frac{2}{3}$" represents a rational number, whereas, "$\frac{\sqrt{2}}{3}$" represents a number which we shall show is not rational. Notice that since the fraction "$\frac{\sqrt{3}}{\sqrt{12}}$" represents the same number as the fraction

"$\frac{1}{2}$", it represents a rational number. The point is that certain fractions name rational numbers and others do not, but all rational numbers can be represented as fractions.

Let us show that the set F of rational numbers includes the integers as a proper subset. Certainly, $\frac{1}{2}$ is a rational number, but not an integer. Thus, $F \neq I$. Furthermore, any integer n may be written as $n = n \cdot \frac{1}{1} = \frac{n \cdot 1}{1} = \frac{n}{1}$, which by definition is a rational number. Thus, $I \subset F$.

We see also that every rational number is a real number. Thus, $F \subset R$. Moreover, if r and s are in F, so are $r + s$, $r - x$, $r \cdot s$ and $\frac{r}{s}$ (if $s \neq 0$), by the results of Chapter 3. Thus, the elements of F satisfy axioms F1 to F7, and $(F, +, \cdot)$ is a sub-field of R. It remains to show that $F \neq R$, that is, there are real numbers which are not rational. This is the problem that now confronts us.

At this point in the development we cannot find any difference between the rational and real number systems; each is an ordered field. But we want the reals to correspond to the whole set of points on the number line, whereas we shall prove shortly that the rationals cannot have this property. The distinction between the reals and the rationals will be made in Chapter 5 by assuming a final property of the real number system.

Since F is an ordered field it must be possible to determine which of two distinct rational numbers is greater.

To do this we need the

Theorem 4.1 If a, b, c, d are real numbers such that $b > 0$ and $d > 0$, then

$$\frac{a}{b} < \frac{c}{d} \longleftrightarrow ad < bc.$$

Proof of $\longrightarrow$:

Since $b > 0$ and $d > 0$, it follows that $bd > 0$. Then

$$\frac{a}{b} < \frac{c}{d} \longrightarrow \left(\frac{a}{b}\right)(bd) < \left(\frac{c}{d}\right)(bd), \qquad \text{O4}$$

$$\longrightarrow \frac{adb}{b} < \frac{bcd}{d}, \qquad \text{Theorem 3.12'}$$

$$\longrightarrow ad < bc, \qquad \text{Theorem 3.17.}$$

The proof of $\longleftarrow$ is left to the reader.

This theorem provides a technique for ordering the rational numbers. Note first that if a, b, c, d are integers, then they are real numbers. Note also that if the rational number $\frac{p}{q}$ has q negative, we can always write it as

$$\left(\frac{p}{q}\right) = \left(\frac{-1}{-1}\right)\left(\frac{p}{q}\right) = \frac{-p}{-q},$$

where now $-q$ is a positive integer. Hence, Theorem 4.1 will apply to any two rational numbers. Now we can determine the relative order of any two rational numbers by comparison with the ordering of two integers. For example,

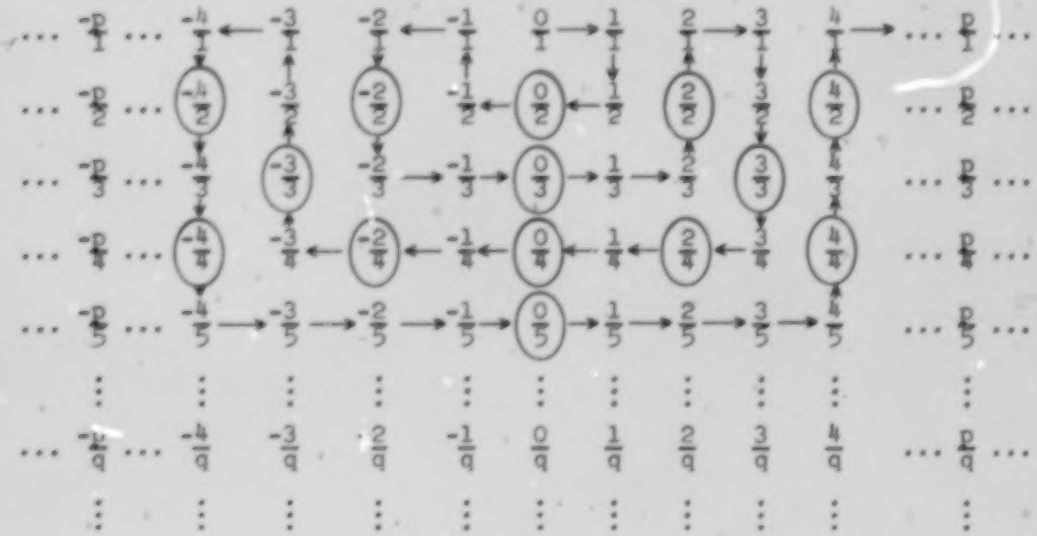
$$\frac{-6}{23} > \frac{-8}{29} \quad \text{because} \quad (-6)(29) > (-8)(23),$$

that is, because $-174 > -184$.

Let us list some properties of the ordered field of rational numbers:

(1) Closure. The set F is closed under addition, subtraction, and multiplication. The quotient of a rational number by a non-zero rational number is a rational number. Thus, with division by zero excluded, F is also closed under division. In the extension from the integers to the rationals we gained a new closure property. Now for any a and b in F there is not only a unique solution of $a + x = b$ in F, but also a unique solution of $ax = b$, $a \neq 0$, in F.

(2) Countability. The set F is countable. To establish this fact let us write the rational numbers in the following array: (Recall that any rational number may be written with a positive denominator.)



Then, moving in the array as indicated by the arrows, starting with $\frac{0}{1}$, we must eventually traverse every rational number whose numerator is an integer and whose denominator is a positive integer. The one-to-one correspondence between N and F is formed as follows, skipping a rational number if it has been encountered previously (circled):

N		F
1	⟷	0
2	⟷	1
3	⟷	$\frac{1}{2}$
4	⟷	$-\frac{1}{2}$
5	⟷	-1
6	⟷	-2
7	⟷	$-\frac{2}{3}$
8	⟷	$-\frac{1}{3}$
9	⟷	$\frac{1}{3}$
⋮		⋮

In this way we are certain that each rational number will correspond to some natural number, and no rational numbers will be overlooked in the process.

(3) Decimal representation. Every rational number r can be represented as a decimal. To illustrate the meaning of this statement let us obtain such a representation of the rational number $\frac{13}{8}$. By the division algorithm,

$$\begin{aligned} 13 &= 8\cdot 1 + 5, \quad 5 < 8, \\ 5\cdot 10 &= 8\cdot 6 + 2, \quad 2 < 8, \\ 2\cdot 10 &= 8\cdot 2 + 4, \quad 4 < 8, \\ 4\cdot 10 &= 8\cdot 5 + 0. \end{aligned}$$

Upon dividing by 8 and successive powers of 10, we obtain

$$\frac{13}{8} = 1 + \frac{5}{8}$$

$$\frac{5}{8} = \frac{6}{10} + \frac{2}{8}\left(\frac{1}{10}\right); \qquad \frac{13}{8} = 1 + \frac{6}{10} + \frac{2}{8}\left(\frac{1}{10}\right)$$

$$\frac{2}{80} = \frac{2}{10^2} + \frac{4}{8}\left(\frac{1}{10^2}\right); \qquad \frac{13}{8} = 1 + \frac{6}{10} + \frac{2}{10^2} + \frac{4}{8}\left(\frac{1}{10^2}\right)$$

$$\frac{4}{800} = \frac{5}{10^3} + 0; \qquad \frac{13}{8} = 1 + \frac{6}{10} + \frac{2}{10^2} + \frac{5}{10^3},$$

and from these equalities the set of inequalities:

$$1 + \frac{6}{10} \leq \frac{13}{8} = 1 + \frac{6}{10} + \frac{2}{8}\left(\frac{1}{10}\right) < 1 + \frac{7}{10}$$

$$1 + \frac{6}{10} + \frac{2}{10^2} \leq \frac{13}{8} = 1 + \frac{6}{10} + \frac{2}{10^2} + \frac{4}{8}\left(\frac{1}{10^2}\right) < 1 + \frac{6}{10} + \frac{3}{10^2}$$

$$1 + \frac{6}{10} + \frac{2}{10^2} + \frac{5}{10^3} = \frac{13}{8}.$$

We abbreviate this to:

$$\frac{13}{8} = 1.625.$$

Notice that for this rational number, one of the remainders in the division algorithm is 0 and the decimal representation therefore terminates; that is, the set of inequalities terminates in an equality. Of course, the whole process can be shortened to the familiar form:

$$\begin{array}{r@{}l}
 & 1.625 \\
8 & \overline{)13.000} \\
 & \underline{8} \\
 & 50 \\
 & \underline{48} \\
 & 20 \\
 & \underline{16} \\
 & 40 \\
 & \underline{40}
\end{array}$$

The reader may wonder why we bothered to write the set of inequalities above, especially since the decimal representation terminated. Another example will show the need for inequalities; let us attempt to represent the rational number $\frac{4}{11}$ as a decimal. Again, by the division algorithm,

$$\begin{array}{ll}
4\cdot 10 = 11\cdot 3 + 7, & 7 < 11, \\
7\cdot 10 = 11\cdot 6 + 4, & 4 < 11, \\
4\cdot 10 = 11\cdot 3 + 7, & 7 < 11, \\
7\cdot 10 = 11\cdot 6 + 4, & 4 < 11, \\
\qquad\vdots & \quad\vdots
\end{array}$$

Dividing by 11 and successive powers of 10, we get

$$\frac{4}{11} = \frac{3}{10} + \frac{7}{11}\left(\frac{1}{10}\right)$$

$$\frac{7}{110} = \frac{6}{10^2} + \frac{4}{11}\left(\frac{1}{10^2}\right); \quad \frac{4}{11} = \frac{3}{10} + \frac{6}{10^2} + \frac{4}{11}\left(\frac{1}{10^2}\right)$$

$$\frac{4}{1100} = \frac{3}{10^3} + \frac{7}{11}\left(\frac{1}{10^3}\right); \quad \frac{4}{11} = \frac{3}{10} + \frac{6}{10^2} + \frac{3}{10^3} + \frac{7}{11}\left(\frac{1}{10^3}\right)$$

$$\frac{7}{11000} = \frac{6}{10^4} + \frac{4}{11}\left(\frac{1}{10^4}\right); \quad \frac{4}{11} = \frac{3}{10} + \frac{6}{10^2} + \frac{3}{10^3} + \frac{6}{10^4} + \frac{4}{11}\left(\frac{1}{10^4}\right)$$

$$\vdots \qquad\qquad \vdots$$

Notice that the remainders repeat in the pattern 7,4,7,4,... and no remainder can be zero. The resulting infinite set of equalities gives rise to the corresponding infinite set of inequalities:

$$\frac{3}{10} \leq \frac{4}{11} < \frac{4}{10}$$

$$\frac{3}{10} + \frac{6}{10^2} \leq \frac{4}{11} < \frac{3}{10} + \frac{7}{10^2}$$

$$\frac{3}{10} + \frac{6}{10^2} + \frac{3}{10^3} \leq \frac{4}{11} < \frac{3}{10} + \frac{6}{10^2} + \frac{4}{10^3}$$

$$\frac{3}{10} + \frac{6}{10^2} + \frac{3}{10^3} + \frac{6}{10^4} \leq \frac{4}{11} < \frac{3}{10} + \frac{6}{10^2} + \frac{3}{10^3} + \frac{7}{10^4}$$

$$\vdots \qquad \vdots$$

When we say that $\frac{4}{11}$ is represented by the infinite decimal .3636... we mean that $\frac{4}{11}$ satisfies every inequality in the above infinite set of inequalities.

The fact that the remainders in the above division algorithm repeat with a fixed pattern leads us to call the resulting decimal representation periodic, and we indicate the set of repeating digits by superscript dots; for example,

$$\frac{4}{11} = .3636\ldots = .\dot{3}\dot{6}.$$

As we have seen, some rational numbers have infinite decimal representations; we shall now show that if the decimal representation of a rational number is infinite, it is also periodic. Consider the positive rational number $\frac{p}{q}$, where p and q are positive integers without common factors. The division algorithm guarantees the existence of integers c and r_0 such that

$$p = qc + r_0, \qquad 0 < r_0 < q.$$

Dividing by q, we have

$$\frac{p}{q} = c + \frac{r_0}{q}, \qquad 0 < \frac{r_0}{q} < 1.$$

Now apply the algorithm to the positive integers $10r_0$ and q:

$$10r_0 = qd_1 + r_1, \qquad 0 \leq r_1 < q,$$

for some integer d_1. Since $r_0 < q$, $10r_0 = qd_1 + r_1 < 10q$ implies that $d_1 < 10$. Again, dividing by $10q$,

$$\frac{r_0}{q} = \frac{d_1}{10} + \frac{r_1}{10q}; \quad \frac{p}{q} = c + \frac{d_1}{10} + \frac{r_1}{10q}.$$

Now if $r_1 = 0$, the decimal representation of $\frac{p}{q}$ terminates. If $r_1 \neq 0$, then $\frac{r_1}{q} < 1$ and

$$c + \frac{d_1}{10} \leq \frac{p}{q} = c + \frac{d_1}{10} + \frac{r_1}{10q} < c + \frac{d_1+1}{10}.$$

Now apply the algorithm to $10r_1$ and q:

$$10r_1 = qd_2 + r_2, \quad 0 \leq r_2 < q,$$

for some integer d_2. Again we can show that $d_2 < 10$ and that

$$\frac{r_1}{10q} = \frac{d_2}{10^2} + \frac{r_2}{10^2q}; \quad \frac{p}{q} = c + \frac{d_1}{10} + \frac{d_2}{10^2} + \frac{r_2}{10^2q}.$$

Now $\frac{r_2}{10q} < 1$ and

$$c + \frac{d_1}{10} + \frac{d_2}{10^2} \leq \frac{p}{q} = c + \frac{d_1}{10} + \frac{d_2}{10^2} + \frac{r_2}{10^2q} < c + \frac{d_1}{10} + \frac{d_2+1}{10^2}.$$

If $r_2 \neq 0$ we continue the process until some remainder r_i is zero or until some remainder r_k is equal to a previous remainder r_j, $j < k$. Thus must occur if no r_i is zero because there are no more than $q - 1$ possible non-zero remainders upon division by In this case the decimal representation of $\frac{p}{q}$ will never terminate and the set of digits $d_jd_{j+1}\ldots d_{k-1}$ will repeat without end.

Thus, if a rational number r has an infinite decimal representation it is periodic, where we mean by a decimal representation of r that for every natural number k, r satisfies the inequality.

$$c + \frac{d_1}{10} + \ldots + \frac{d_k}{10^k} \leq r < c + \frac{d_1}{10} + \ldots + \frac{d_k+1}{10^k},$$

where each d_i is some integer in the set

$$\{0,1,2,\ldots,9\}.$$

Again we should point out that the digits of the representation of r may be restricted to any set of the form

$$\{0, 1, \ldots, (b-1)\}, \quad b > 1.$$

The development of the representation given above for $b = 10$ is quite general and does not depend in any way on the scale or base b of representation. For example, the rational number $\frac{4}{3}$ may be represented in the "five" scale as follows:

$$\begin{array}{lllll}
4 = 3\cdot 1 + 1 & \Longrightarrow \frac{4}{3} = 1 + \frac{1}{3} & & \Longrightarrow 1 \leq \frac{4}{3} < 2 & \\
5\cdot 1 = 3\cdot 1 + 2 & \Longrightarrow \frac{1}{3} = \frac{1}{5} + \frac{2}{5\cdot 3} & & \Longrightarrow 1 + \frac{1}{5} \leq \frac{4}{3} < 1 + \frac{2}{5} & \\
5\cdot 2 = 3\cdot 3 + 1 & \Longrightarrow \frac{2}{3} = \frac{3}{5} + \frac{1}{5\cdot 3} & & \Longrightarrow 1 + \frac{1}{5} + \frac{3}{5^2} \leq \frac{4}{3} < 1 + \frac{1}{5} + \frac{4}{5^2} & \\
\vdots & \vdots \quad \vdots & & \vdots & \vdots \quad \vdots
\end{array}$$

Hence,

$$\frac{4}{3} = 1.1313\ldots_{\text{five}} = 1.\dot{1}\dot{3}_{\text{five}}.$$

It can also be shown (see Appendix A) that every periodic decimal represents a rational number and only one rational number.

(4) Density. The set of rational numbers is dense; that is, for any two rational numbers u and v, there is a rational number w such that $u < w < v$. Viewed on the number line, this property asserts that any two points with rational number coordinates, no matter how

close together, have a point between them with rational coordinate. But there is no end to this argument; this implies that between any two rational numbers there are infinitely many rational numbers. The proof of this property follows immediately from Problem 5 on page 3.24. If a and b are rational numbers such that $a < b$, then

$$a < \frac{a+b}{2} < b.$$

Since $\frac{a+b}{2}$ is also rational if a and b are rational, the density property is established.

At first thought, one would suspect that the set F, being dense, corresponds to the set of all points on the number line. There would seem to be no "room" between the rationals. This is not the case. When we introduce a final axiom for the real number system it will be possible to prove, for example, that there is a positive real number x such that $x^2 = 2$, that is, $\sqrt{x} = 2$. We shall now prove that there is no rational number x such that $x^2 = 2$.

Theorem 4.2 There is no rational number x such that $x^2 = 2$; that is, there do not exist two integers a and b without common factors such that $x^2 = (\frac{a}{b})^2 = 2$.

Proof: We shall use a proof by contradiction. Assume as part of the hypothesis that there are two integers a and b, without common factors, such that $(\frac{a}{b})^2 = 2$. We underline the restriction "without common factors" because by Theorem 3.17 it is always possible to reduce a rational number $\frac{a}{b}$ to such a form. Then

$$(\frac{a}{b})^2 = 2 \Longrightarrow a^2 = 2b^2.$$

Since b is an integer, so is b^2. Thus a^2 is an even integer. By the result of Problem 1(c) on page 4.15, a is also an even integer; thus, we may write

$$a = 2c, \quad \text{for some integer } c.$$

Then

$$a = 2c \Longrightarrow a^2 = 4c^2.$$

We now have $a^2 = 4c^2$ and $a^2 = 2b^2$, so that $2b^2 = 4c^2$; that is, $b^2 = 2c^2$. But c^2 is an integer, so that b^2 is an even integer. This means that b is also an even integer. We have arrived at a contradiction, for if a and b are both even integers they must have the common factor 2, contrary to our original requirement. Thus, there is no rational number x such that $x^2 = 2$.

Real numbers that are not rational are called irrational numbers. Thus, $\sqrt{2}$ is an irrational number.

Let us summarize what has been found. In the transition from the integers to the rational numbers, we lost some properties. For one, we do not have unique factorization of rational numbers. If r and s are any non-zero rational numbers, then r is a factor of s and s is a factor of r; that is, there exist rational numbers u and v such that $s = ru$ and $r = sv$. Another property we lost is that of terminating decimal representation. But we gained the important properties of closure under division (0 excluded) and density.

This concludes the discussion of three proper sub-systems of the reals:

$$N \subset I \subset F \subset R.$$

Which properties of F are shared by R? Certainly R has the same closure properties as F. But it will be shown (see Appendix A) that R is not countable. R is also dense, and its elements can be represented in decimal form, although these decimals will be shown to be non-periodic in general.

What new properties does R have? It will be shown that the set of positive real numbers is closed under the extraction of a root. This means, for example, that for a in R the equation $x^2 = a$ has a solution in R if $a > 0$. The most important new property, from the standpoint of analysis and geometry, is the fact that R is complete. This will be the theme of the next chapter.

Exercises

1. Determine the relative order of the rational numbers

$$\frac{-37}{61},\quad \frac{4}{5},\quad \frac{-12}{20},\quad \frac{47}{59}.$$

2. The rational number $(\frac{3}{16})_{ten}$ can be represented as the decimal $\frac{1}{10} + \frac{9}{10^2} + \frac{7}{10^3} + \frac{5}{10^4} = (.1875)_{ten}$.

   (a) Write the number $(.302)_{four}$ in the form of a rational number in base ten.

   (b) Convert $(1.5)_{ten}$ to a decimal in the four scale.

3. Prove: There is no rational number x such that $x^2 = 3$.

4. Prove: There is no rational number x such that $x^3 = 2$.

5. Show that if n is a natural number, then $0 < \frac{1}{n} \leq 1$.

6. Show that if n is a natural number, then $0 < \frac{1}{n^2} \leq \frac{1}{n}$.

7. Consider the set T of positive rational numbers, with an ordering given by the axioms O1 to O4.

   (a) Prove that T is not well ordered; that is, there is a subset of T which does not have a least element.

   (b) Are there rational numbers which are less than every element of T? Which of these is greatest?

8. Using the division algorithm, develop the decimal representation of $\frac{1}{7}$.

9. Find the truth set of the sentence

$$(x - 1)(x + 1)(2x - 3)(x^2 - 2) = 0$$

if the domain of x is the

(a) set of natural numbers,

(b) set of integers,

(c) set of rational numbers,

(d) set of real numbers.

10. Factor the polynomial $x^4 - 9$ into polynomials with coefficients

(a) in F, (b) in R.

11. For the elements of the set of positive rational numbers, let us define a prime p to be a number that cannot be obtained as the product of numbers in the set that are all less than p. For example, $\frac{1}{3}$ is prime because if $ab = \frac{1}{3}$ and a, b are positive rational numbers, then either a or b is greater than $\frac{1}{3}$. On the other hand, $\frac{4}{3}$ is composite, since $\frac{4}{3} = \frac{7}{6} \cdot \frac{8}{7}$ and $\frac{7}{6}$, $\frac{8}{7}$ are each less than $\frac{4}{3}$. With this definition of a prime, show by a counter-example that the positive rationals do not have the property of unique factorization.

Chapter 5

COMPLETENESS OF THE REAL NUMBER SYSTEM

1. The Completeness Axiom. Our attention thus far has been on the axiomatic development of an abstract system whose model is the real number system. Before we continue this development and bring it to a satisfactory conclusion, it will help to review some geometric ideas. In geometry it is assumed that the points of a line satisfy a certain set of geometric axioms.* One of these states that a line contains at least two distinct points. Another assumes that to every pair of points A, B there corresponds a unique real number called the measure of the distance between A and B. A third axiom says that given two different points A and B on a line L, there is a one-to-one correspondence between the points of L and the real numbers such that

(1) A corresponds to zero;

(2) B corresponds to a positive number;

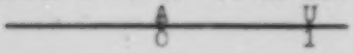
(3) if P and Q are any points on L and if P corresponds to x and Q to y, then the measure of the distance between P and Q is $|y - x|$.

With A corresponding to zero, let us mark the point U to the right of A so that the measure of the distance between U and A is 1;

---

*See Volume II of this series of Studies in Mathematics: Euclidean Geometry Based on Ruler and Protractor Axioms, by Curtis, Daus, and Walker.

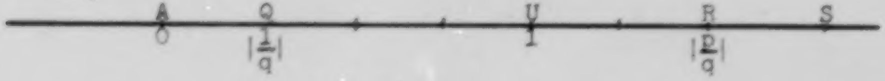
then assign the number 1 to U. It is now possible to



prove the theorem*: If three distinct points X, Y, Z of the line have coordinates x, y, z, respectively, then Y is between X and Z if and only if $x < y < z$ or $z < y < x$.

It is now apparent how (and why) we may form the unique correspondence between the rational numbers and points of the number line. The rational number $\frac{p}{q}$, p and q in I and $q \neq 0$, is assigned to a point as follows: Determine the point R to the right of A such that the measure of the distance between R and A is $|\frac{p}{q}|$. Then assign the positive rational number $|\frac{p}{q}|$ to the point R and the negative rational number $-|\frac{p}{q}|$ to the point R' on the left of A and at the same distance from A.

The construction of the point R on the line is accomplished as follows: By dividing the segment AU into $|q|$ equal parts, determine the point Q such that the measure of the distance between A and Q is $|\frac{1}{q}|$. (The reader should recall how this is done with straight edge and compass.) Then lay off $|p|$ of these distances from A to the right, terminating at a point R. The distance between A and R then has measure $|\frac{p}{q}|$, and we assign the positive rational number $|\frac{p}{q}|$ to R.



It follows from the stated geometric axioms and theorem and the orientation of the rationals on the line that if r and s are the coordinates of R and S, respectively, then

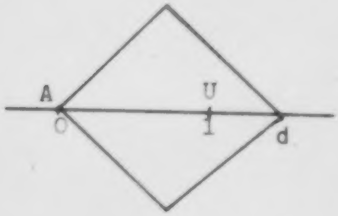
---

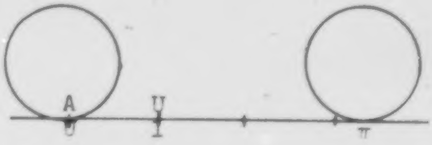
*Theorem 4.1 on page 4.31 of Volume II, Studies in Mathematics.

$$r < s \Longleftrightarrow R \text{ is to the left of } S.$$

To prove $\Longrightarrow$ we consider the three cases: (1) $r \geq 0$ and $s \geq 0$, (2) $r < 0$ and $s < 0$, (3) $r < 0$ and $s \geq 0$. In case (1) R and S are both to the right of A. Then, since $0 \leq r < s$, it follows that either R is between A and S or R coincides with A. In either case, R is to the left of S. In case (2) $r < s < 0$ implies that R and S are both to the left of A, and S is between R and A. Hence, R is to the left of S. In case (3) $r < 0 \leq s$ implies that R is to the left of A and S is either to the right of A or coincides with A. In either case, R is to the left of S. The arguments are easily revised to prove $\Longleftarrow$.

The Greek mathematician Pythagoras reasoned that the length d of the diagonal of a square with side of length 1 satsifies the equation $d^2 = 2$. He concluded that there is a "number" d satisfying this equation because d measures a length. On the basis of our geometric axioms there is a real number d measuring the distance between the opposite vertices of the unit square.





In the same way we find a real number, which we call π, which measures the distance between the beginning and ending points of tangency as a circle of unit diameter rolls through one revolution on a line. But Pythagoras believed that all numbers are rational, the ratios of integers. When he finally proved that there is no rational number whose square is 2 (in essentially the same way it was proved in Theorem 4.2) he found himself in a dilemma. Later, the number π was also found to be not rational.

It is precisely this same dilemma that faces us at this point. We have assumed the existence of a set of elements called real numbers which satisfy the axioms of an ordered field. But a proper subsystem of the reals, the system of rational numbers, satisfies the same axioms and apparently has the same properties as the reals. We have established that the rationals do not include numbers such as x, where $x^2 = 2$. On the other hand we can find a point on the number line the square of whose distance from 0 is 2. That is, there are points on the number line to which no rational numbers can be assigned.

Our objective, then, is to complete the description of the system of real numbers in such a way that to each point on the number line there will be assigned exactly one real number.

Before stating the axiom which will complete the description, we need the

Definition. A non-empty set S of real numbers is *bounded above* if there exists a real number M such that $s \leq M$ for every s in S. The number M is called an *upper bound* of S. A real number L is a *least upper bound* (lub) of S if

(1) L is an upper bound of S, and

(2) for every upper bound M of S, $L \leq M$.

For example, the finite set $T = \{1,3,5,8,17\}$ has 18 as an upper bound. In fact, any number 17 or greater serves as an upper bound of T. Obviously, the lub of T is 17. The infinite set

$$U = \{1/3, 2/5, 3/7, 4/9, \ldots, \frac{n}{2n+1}, \ldots\}$$

has 1/2 as an upper bound, since

$$\frac{n}{2n+1} < \frac{1}{2} \text{ for every } n \text{ in } N. \text{ (Prove this.)}$$

It can also be shown that no real number c less than 1/2 is an upper bound of U. Hence, the lub of U is 1/2. In this case the lub of the set is not an element of the set.

If we restrict our attention to the set F of rational numbers, the question arises: Does every bounded, non-empty set in F have a least upper bound in F? Consider, for example, the set S of all positive rational numbers s such that $s^2 < 2$. This set is not empty (since 1 is in S) and it has an upper bound 2 in F (since $s^2 < 2$ and $2 < 2^2 \Longrightarrow s^2 < 2^2 \Longrightarrow s < 2$). It will turn out that this bounded set in F does not have a least upper bound in F.

Here we have the basic difference between the rationals and the reals. It is stated as our final axiom.

C (Completeness axiom) Every non-empty set of real numbers which has an upper bound in R has a least upper bound in R.

The least upper bound guaranteed by this axiom is unique. To show this, assume the non-empty set S of real numbers is bounded above and has two least upper bounds L and L'. Then both L and L' are upper bounds of S. By definition, $L' \leq L$, since L is a lub of S. For the same reason, $L \leq L'$. Hence, $L = L'$, and the lub of S is unique.

With axiom C (sometimes called the continuity axiom) we shall be able to prove that, for example, $\sqrt{2}$ is a real number. In other words, there is a real number x such that $x^2 = 2$. Furthermore, we can show more generally that the equation $x^2 = a$, $a > 0$, has a solution in R.

Exercises

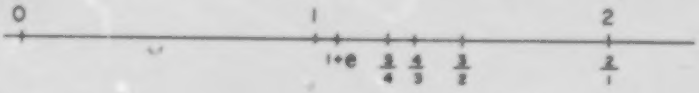
1. Give two upper bounds and the lub of each of the following sets:
   (a) $\{\frac{1}{2}, \frac{2}{3}, -\frac{3}{4}, \frac{5}{7}\}$

   (b) $\{-3.6, -3.62, -3.615, -3.654\}$
2. Let T be the set of all real numbers t less than 1. Prove that 1 is the lub of T. (Assume some real number c, $c < 1$, is an upper bound of T. Then apply the inequality $a < \frac{a+b}{2} < b$ to c and 1, and show there is a real number in T which is greater than c.)
3. Write a corresponding definition of a lower bound of a non-empty set and greatest lower bound (glb) of the set.
4. Prove: Every non-empty set of real numbers which has a lower bound in R has a greatest lower bound in R. (Let s be any element of a non-empty set S with lower bound m; then
   $s \geq m \Longrightarrow -s \leq -m$. Hence, the set S' of all opposites -s of elements of S has upper bound -m. Apply axiom C to S' to obtain the lub of S', say -L; then show that L must be the glb of S.)
5. Find upper and lower bounds of the sets:
   (a) $\{\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \dots, \frac{n}{n+1}, \dots\}$
   (b) $\{\frac{2}{1}, \frac{3}{2}, \frac{4}{3}, \dots, \frac{n+1}{n}, \dots\}$
   (c) $\{\frac{1}{2}, \frac{2}{5}, \frac{3}{10}, \frac{4}{17}, \dots, \frac{n}{n^2+1}, \dots\}$

(d) $\{\frac{2}{1}, \frac{5}{4}, \frac{10}{9}, \frac{17}{16}, \ldots, \frac{n^2+1}{n^2}, \ldots\}$

6. Find the lub of the set (a) and the glb of each of (b), (c), (d) in Problem 5. Try to prove these results.

   For example, the glb of the set $\{\frac{2}{1}, \frac{3}{2}, \frac{4}{3}, \ldots, \frac{n+1}{n}, \ldots\}$ is 1. To prove this, we show that 1 is a lower bound of the set and that every number greater than 1 is not a lower bound.



   For every natural number n, $n + 1 > n$ (why?) and, hence,

$$\frac{n+1}{n} > 1.$$

   That is, 1 is a lower bound of the set. Consider any number greater than 1, say $1 + e$, where e is an arbitrarily small positive real number. By means of the computation

$$\frac{n+1}{n} - 1 = \frac{1}{n} < e \Longleftrightarrow n > \frac{1}{e},$$

   we see that by choosing a natural number n such that $n > \frac{1}{e}$,

$$n > \frac{1}{e} \Longrightarrow \frac{1}{n} < e \Longrightarrow \frac{n+1}{n} - 1 < e \Longrightarrow \frac{n+1}{n} < 1 + e;$$

   that is, there is an element of the set that is less than $1 + e$. Hence, 1 is the glb of the set. It should be pointed out that we assumed intuitively the existence of a natural number greater than $\frac{1}{e}$. We shall prove this in the next section.

7. Prove that if an element of a set is an upper bound of the set, it must be the lub of the set.

2. The Existence of $\sqrt{2}$ in R. We know that the equation $x^2 = 2$ has no solution in F. It is our purpose in this section to show how the completeness axiom guarantees that $x^2 = 2$ does have a solution in R. An outline of the proof is as follows:

> Consider the set T of all positive real numbers t such that $t^2 < 2$, and show that T has a lub, say x. Then exactly one of the following must be true:
>
> $$x^2 < 2, \quad x^2 > 2, \quad x^2 = 2.$$
>
> If we can show that the first two cannot be true, we have the conclusion, $x^2 = 2$.

In preparation for the proof we need three supporting theorems. The first says that for any two positive real numbers there can be found a natural number multiple of one number which is greater than the other. This result is then used to prove the second and third. The second states that for any positive real number whose square is less than 2 there is a greater real number whose square is also less than 2. The third asserts that for any positive real number whose square is greater than 2 there is a lesser positive real number whose square is also greater than 2. A moment of reflection will show that the second and third theorems, when proved, will rule out the possibilities of $x^2 < 2$ or $x^2 > 2$ if x is the lub of the set of positive real numbers whose squares are less than 2.

Theorem 5.1 For every two positive real numbers a and b there is a natural number n such that $na > b$.

Proof by contradiction: For a given pair of positive real numbers a and b assume there is no n in N such that $na > b$. Then the set U of all products of the form na has the property that $na \leq b$ for every n in N. Hence, b is an upper bound of U. Now since U is non-empty, we know by axiom C that U has a lub in R, say c, such that every element of U is less than or equal to c. Since $n + 1$ is in N if n is in N, the number $(n + 1)a$ is in U. Then

$$(n + 1)a \leq c$$

$$na + a \leq c$$

$$na \leq c - a, \quad \text{for every } n \text{ in N.}$$

Thus, $c - a$ is an upper bound of U. But $c - a < c$, since $a > 0$. Here we have a contradiction, for we cannot have an upper bound $c - a$ less than the lub c. Hence, the theorem is proved.

An ordered field which has the property of this theorem is called Archimedean. Both $(R, +, \cdot)$ and $(F, +, \cdot)$ are Archimedean.

We use this result to prove

Lemma 5.2 If a is any positive real number such that $a^2 < 2$, then there exists a real number b such that $b > a$ and $b^2 < 2$.

Proof: Given a in R, $a > 0$, $a^2 < 2$, let us construct the real number

$$b = a + \frac{a}{n},$$

where n is in N, and show that for some n

$$b > a \text{ and } b^2 < 2.$$

For any n in N, $a + \frac{a}{n} > a$ (why?); hence, $b > a$ for every n in N.

Now we compute, for any n in N,

$$b^2 = (a + \frac{a}{n})^2 = a^2(1 + \frac{1}{n})^2 \leq a^2(1 + \frac{3}{n}).$$

(Here the reader should pause and verify for himself that

$$(1 + \frac{1}{n})^2 \leq 1 + \frac{3}{n} \text{ for every } n \text{ in N.})$$

On the other hand,

$$a^2 < 2 \longrightarrow \frac{2}{a^2} > 1 \longrightarrow \frac{2}{a^2} - 1 > 0.$$

Since $\frac{2}{a^2} - 1$ is positive, so is its reciprocal, and by Theorem 5.1 there is some n in N such that

$$n(\frac{1}{3}) > \frac{1}{\frac{2}{a^2} - 1}.$$

Then,

$$\frac{3}{n} < \frac{2}{a^2} - 1 \quad \text{(why?)}$$

$$\frac{3}{n} + 1 < \frac{2}{a^2}$$

$$a^2(1 + \frac{3}{n}) < 2 \text{ for some } n \text{ in N.}$$

Putting these inequalities together, for some n in N,

$$b^2 \leq a^2(1 + \frac{3}{n}) \text{ and } a^2(1 + \frac{3}{n}) < 2 \longrightarrow b^2 < 2.$$

This concludes the proof.

Lemma 5.3 If a is any positive real number such that $a^2 > 2$, then there exists a positive real number b such that $b < a$ and $b^2 > 2$.

Proof: The proof follows that of Lemma 5.2. Construct the real number

$$b = a - \frac{a}{n},$$

where n is in N, and show that for some n,

$$b > 0, \quad b < a, \quad b^2 > 2.$$

The reader can show that $b > 0$ and $b < a$ for any n in N, $n > 1$. Now compute

$$b^2 = (a - \frac{a}{n})^2 = a^2(1 - \frac{1}{n})^2 \geq a^2(1 - \frac{2}{n}).$$

Also,

$$a^2 > 2 \longrightarrow \frac{2}{a^2} < 1 \longrightarrow 1 - \frac{2}{a^2} > 0.$$

Then by Theorem 5.1 there is an n in N such that

$$n(\frac{1}{2}) > \frac{1}{1 - \frac{2}{a^2}}$$

$$\frac{2}{n} < 1 - \frac{2}{a^2}$$

$$a^2(1 - \frac{2}{n}) > 2 \text{ for some } n \text{ in N.}$$

Then by transitivity of inequalities, for some n in N

$$b^2 \geq a^2 (1 - \frac{2}{n}) \text{ and } a^2(1 - \frac{2}{n}) > 2 \longrightarrow b^2 > 2,$$

and the theorem is proved.

The stage is now set for the main theorem of the section.

Theorem 5.4 (Existence of $\sqrt{2}$ in R) There is a positive real number x such that $x^2 = 2$.

Proof: Consider the set T of all positive real numbers t such that $t^2 < 2$. Certainly 1 is in T, since $1^2 < 2$; also

$$t^2 < 2 \text{ and } 2 < 2^2 \longrightarrow t^2 < 2^2 \longrightarrow t < 2,$$

and we see that 2 is an upper bound of T. Thus by axiom C the set T has a lub, say x. Now by O1 we are assured that exactly one of the following sentences is true:

$$x^2 < 2, \quad x^2 > 2, \quad x^2 = 2.$$

We shall rule out the first two as follows:

(1) Let x be the lub of T and assume that $x^2 < 2$. Then Lemma 5.2 asserts that there is a positive real number b such that

$$b > x \text{ and } b^2 < 2.$$

Thus, b is in T (since $b^2 < 2$). But this is a contradiction, for we cannot have any element of a set greater than its lub. Hence, "$x^2 < 2$" is false.

(2) Let x be the lub of T and assume that $x^2 > 2$. Then Lemma 5.3 exhibits a positive real number b such that

$$b < x \text{ and } b^2 > 2.$$

Now for any element t of T, $t^2 < 2$, so that

$$t^2 < 2 \text{ and } 2 < b^2 \Longrightarrow t^2 < b^2 \Longrightarrow t < b.$$

Thus, b is an upper bound of T. This is a contradiction, for b is less than the lub x. Hence, "$x^2 > 2$" is false, and we have remaining only the sentence "$x^2 = 2$". This proves the theorem.

The reader may think that a great deal of effort has gone into a simple result. On the contrary, we have opened a vast domain of new numbers. It is now a simple matter to reword Lemmas 5.2, 5.3, and Theorem 5.4 by replacing the number 2 by any positive real number c. The result is

*Theorem 5.5 If c is any positive real number, then there is a unique positive real number x such that $x^2 = c$.

It is more tedious, but possible, to go even farther and prove that if c is any positive real number and n is any natural number, then there exists a positive real number x such that $x^n = c$.

Definition. If a is a positive real number, the unique positive solution of $x^2 = a$ is called the square root of a and denoted by $\sqrt{a}$. The other solution of $x^2 = a$ is therefore $-\sqrt{a}$. In general, if a is a positive real number, the unique positive solution of $x^n = a$, n in N, is called the nth root of a and denoted by $\sqrt[n]{a}$. ($\sqrt[n]{0} = 0$)

As a consequence of the above definition we have

$$\sqrt{a^2} = |a|.$$

For example, $\sqrt{(-3)^2} = |-3| = 3$ and $\sqrt{(x - 1)^2} = |x - 1|$, where the absolute value notation guarantees that the square root is positive (or zero).

Real numbers that are not rational are called irrational. Thus, $\sqrt{2}$ is an example of an irrational number. But not all irrational numbers are of the form $\sqrt[n]{a}$. Other real numbers such as π, which are not solutions of polynomial equations, are irrational. (See Appendix C.) Our task is not completed with the proof of the existence of $\sqrt[n]{a}$, $a \geq 0$, in R. We still need to show that to every point on the number line there corresponds a real number.

Exercises

1. Prove that there is a positive real number x such that $x^2 = 3$.
2. Show that there is a positive rational number c such that $c > \frac{141}{100}$ and $c^2 < 2$.
3. Show that there is a positive rational number d such that $d < \frac{142}{100}$ and $d^2 > 2$.
4. Prove: If a is in F and b is irrational, then
   (a) $a + b$ is irrational,
   (b) ab is irrational,

(c) $\frac{1}{b}$ is irrational.

5. Prove: The equation $x^2 + bx + c = 0$, with b and c in R, has a solution in R if and only if $b^2 - 4c \geq 0$; there are one or two distinct solutions according as $b^2 - 4c = 0$ or $b^2 - 4c > 0$.

6. Prove the corollary to Theorem 5.1: For any two positive real numbers a and b there is a natural number n such that

$$n > \frac{b}{a} \quad \text{and} \quad \frac{b}{n} < a.$$

7. Use the corollary in Problem 6 to prove the existence of a natural number n such that $n > \frac{1}{e}$ for any given positive e.

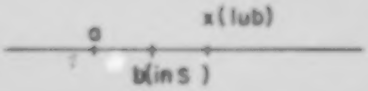
---

3. Completeness of the Set of Reals. In the previous section we proved that certain numbers, such as $\sqrt{2}$, are in R even though not in F. It remains to show that every point on the number line can be assigned a real number coordinate in only one way.

Again, we need three preparatory theorems, each of which is used in the proof of the succeeding theorem.

Lemma 5.6 Given a non-empty set S of real numbers with lub x in R, if a is any number in R such that $a < x$, then there is a number b in S such that $b > a$.

Discussion: It is instructive to view this theorem on the number line:



It says that for any real number a less than the lub x of S (no matter how close a is to x) there is a number in S which is between a and x.

Proof by contradiction: Assume there is no element of S greater than a. Then every element of S is less than or equal to a, and a is an upper bound of S. This is a contradiction, for $a < x$ and x is the lub of S. Hence, there is an element, say b, of S greater than a.

This lemma is used to prove

Lemma 5.7 For a given number a in R let S be the set of all rational numbers x such that $x < a$. Then a is the lub of S.

Proof: Since $x < a$ for all x in S, a is an upper bound of S. By axiom C we know that S has a lub, say y, such that $y \leq a$. If we can show that $y \nless a$, then $y = a$ and the theorem is proved.

Assume that $y < a$, that is, $a - y > 0$. Then by Theorem 5.1 there is an n in N such that

$$n(1) > \frac{1}{a - y}$$

$$\frac{1}{n} < a - y$$

$$y + \frac{1}{n} < a.$$

Also, $y - \frac{1}{n} < y$. Now by Lemma 5.6 there is a z in S such that

$$y - \frac{1}{n} < z \text{ and } z \leq y.$$

Then

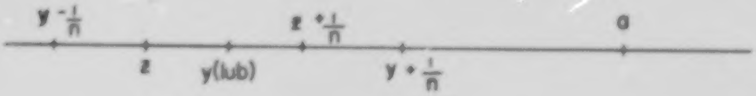
$$y - \frac{1}{n} < z \leq y \longrightarrow y < z + \frac{1}{n} \leq y + \frac{1}{n} < a,$$

and

$$y < z + \frac{1}{n} < a.$$

This is a contradiction, because $z + \frac{1}{n}$ is in S (it is a rational number less than a) and cannot be greater than y, the lub of S. Hence, "$y < a$" is false, and we have proved that $y = a$.

Notice how the theorem looks on the number line.



If y is less than a, then there is some number $z + \frac{1}{n}$ in S such that $z + \frac{1}{n} > y$.

The final preparatory theorem states that between any two distinct real numbers there lies a rational number. In other words, the rational numbers are dense in R.

Theorem 5.8 If a and b are any two distinct real numbers such that $a < b$, then there is a rational number c such that $a < c < b$.

Proof: Let S be the set of all rational numbers x such that $x < b$. Then by the previous lemma, b is the lub of S. Now by Lemma 5.6

$$a < b \Longrightarrow c > a \text{ for some } c \text{ in } S.$$

But if c is in S, then $c < b$. Hence,

$$a < c < b,$$

where c is a rational number.

Now we can prove the main theorem.

Theorem 5.9 Corresponding to each point P of the number line there is exactly one real number x.

Proof: Let P be a point on the number line to which no rational number has been assigned. We shall show that there is a unique real number x corresponding to P.

Let L be the set of all rational numbers corresponding to points to the left of P, and let R be the set of all rational numbers corresponding to points to the right of P. Now each rational number is in either L or R but not both.

Let a be the lub of L and b the glb of R. Then either

$$a < b \text{ or } a = b.$$

Assume that $a < b$. Then by Theorem 5.8 there is a rational number c such that $a < c < b$. Hence c, being rational, is in L or R, but not both. This is a contradiction, for c cannot be in L (being greater than a) or in R (being less than b). Thus, $a \not< b$, and we have shown that a = b. Since the lub of a set is unique, there is exactly one real number x = a = b corresponding to P.

With this theorem we have finished the description of the system of real numbers as a complete ordered field. The consequences are far-reaching in many fields of mathematics, such as analytic geometry, calculus, numerical analysis, to name only a few.

But one of our primary objectives is still not attained; we have no assurance that the equation $x^2 = a$, a in R, has a solution in R. (Only if $a \geq 0$ does it have a solution in R.) When one thinks of the great variety of algebraic equations that may be encountered, it is questioned whether we can ever develop a number system adequate to provide solutions for all algebraic equations. Fortunately, only one more extension is necessary, an extension to the complex number system, to guarantee all such solutions.

A first course is usually not concerned with complex numbers; thus the extension to the complex number system and a discussion of the properties of these numbers is deferred to Appendix B.

How may real numbers be represented? One of the consequences of the theorems of this section is the fact (see problem 4) that a given irrational number may be approximated as closely as desired by a rational number. A discussion of representation is given in Appendix A.

How much of the theory of this chapter should be included in a first course? Very little. The purpose of the chapter was to provide a clear understanding in the teacher's mind of the nature and character of the real number system. Only then can he transmit a correct intuitive picture of real numbers to his students.

Exercises

1. Prove the counterpart of Lemma 5.6: Given a non-empty set T of real numbers with glb y in R. If b is any real number such that $b > y$, then there is a number c in T such that $c < b$.
2. Prove the counterpart of Lemma 5.7: For a given b in R let T be the set of all rational numbers y such that $y > b$. Then b is the glb of T.
3. Prove the statement made in Section 1: The set S of all rational numbers s such that $s^2 < 2$ does not have a lub in F. (Between the real numbers a and $\sqrt{2}$, $a < \sqrt{2}$, there is a rational number, by Theorem 5.8.)
4. Prove: If ε is any given arbitrarily small positive real number and a is any given real number, then there is a rational number x such that $|x - a| < \varepsilon$. (That is, any real number can be approximated as closely as desired by a rational number. Since $a - \varepsilon < a + \varepsilon$ then between $a - \varepsilon$ and $a + \varepsilon$ there is a rational number, by Theorem 5.8.)
5. Let I_n be the set of numbers x satisfying

$$a_n \leq x \leq b_n, \quad a_n < b_n.$$

If $a_n \leq a_{n+1}$, $b_{n+1} \leq b_n$, and if $b_n - a_n = \frac{1}{10^n}$, then the set $\{I_1, I_2, \ldots, I_n, \ldots\}$ is called a *nest of intervals*. Prove that under these conditions there is *exactly one* real number which is in *every* I_n. (Show that the set $\{a_1, a_2, \ldots, a_n, \ldots\}$ is bounded above; hence, has a unique lub a. Then show that a is in every I_n. Finally, show that if c is in every I_n and $c \neq a$, there is a contradiction.)

Chapter 6

FUNCTIONS

1. Variables. In our discussion of the real number system we wrote statements such as:

(1) For any a and b in R, $a + b = b + a$.

In other contexts in algebra we see statements such as the formal equation

$$\frac{x^2 - 4}{x + 2} = x - 2. \qquad (2)$$

In these sentences there occur symbols $x, y, z, a, b, c, \ldots$ and numbers from some set of numbers. We usually refer to such symbols as variables, using the word loosely. But it is not clear that variables play the same role or have the same meaning in each of the sentences (1) and (2). In (1) a and b represent any elements in R; in this context a variable is a quantified symbol representing an element of a given set. **The quantifier in (1) is "any".** In the sentence, "The integer a is even if there is some integer b such that $a = 2b$," the quantifier is "some", meaning "at least one."

On the other hand there are no quantifiers in (2). Here the symbol x has not been restricted in any way. In this context a variable is an indeterminate, an unquantified symbol which has no meaning until it is agreed what properties it enjoys. Indeterminates were used in this way for a long time before anyone succeeded in clarifying their significance.

Are these two meanings of "variable" compatible? How may these meanings be used interchangeably in algebra without confusion? And most important, how may all the concepts of variables, operations, relations and correspondences be unified in algebra? These questions will be considered in the following sections.

2. Algebraic Expressions. We devoted much of our consideration in this study to the algebra of the real number system, in which "variable" is used in the sense of a "quantified symbol" representing an element of R or of one of the subsets of R. Now let us develop, if possible, an algebra of indeterminates. Then we shall show how these two meanings of "variable" complement and abet each other in much of the first course in algebra.

First, we fix our attention on a set S of numbers and the set of four binary connectives $+, -, \times, \div$. The set S can be any subset of R, or possibly the set of complex numbers (see Appendix B). The connectives, when applied to pairs of numbers in S, are the usual field operations. Now let us attach to the set S any indeterminate symbols, say x, y, z, a, b, c, ... From this enlarged set we build algebraic expressions over the set S by the following

Definition

(1) Each of the symbols and numbers is an algebraic expression.

(2) Given any two algebraic expressions A and B and a connective * from the set of connectives, then A*B is an algebraic expression.

(3) If A is an algebraic expression, then $\sqrt[n]{A}$ is an algebraic expression, where n is in N.

(4) Any finite number of operations (1), (2) or (3) results in an algebraic expression.

Thus, algebraic expressions over S are those which can be constructed from indeterminates and numbers in S by the above rules of formation, somewhat in the same way that English expressions are formed from words by certain rules of grammar. For example,

$$\frac{\sqrt[4]{x - y}}{2x + 3}$$

is an algebraic expression* over I. (The reader can verify this by tracing the sequence of operations which generates the expression from the symbols x, y and numbers from I.)

On the other hand,

$$\frac{\sqrt[+]{x -}}{+ 3}$$

is not an algebraic expression because it does not conform to prescribed mathematical "grammar"; whereas

$$\sin x$$

is not an algebraic expression because it requires an infinite sequence of permissible operations for its representation:

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$$

Now consider the system of all algebraic expressions over the set R of all real numbers and the binary operations of addition and multiplication. By definition, corresponding to each element A of this system there is an element (-1)A and an element 1 ÷ A in the system. Let us agree to abbreviate "(-1)A" to "-A" and "1 ÷ A" to "$\frac{1}{A}$".

---

*We agree that "A ÷ B" may be written "$\frac{A}{B}$".

With "A = B" meaning "A is the same algebraic expression as B", let us define subtraction and division by :

$$A - B = A + (-B) \quad \text{and} \quad A \div B = A(\tfrac{1}{B}) .$$

Here we are again in a familiar position. We have a system of elements and two binary operations such that for any two elements A and B of the system, A + B and AB are unique elements of the system. But certain elements of this system (the real numbers) already have their properties prescribed under the operations, so that we are not free to assume properties of algebraic expressions arbitrarily. Instead we shall extend the definition of equality of algebraic expressions in such a way that the field axioms F1 to F7 hold. That is, if A, B, C are any algebraic expressions, we define A + B and B + A to be equal, AB and BA to be equal, (A + B) + C and A + (B + C) to be equal, etc. It can be shown that this definition of "=" for algebraic expressions has the desired equivalence properties:

$$A = A; \quad A = B \Longleftrightarrow B = A; \quad A = B \text{ and } B = C \Longrightarrow A = C.$$

The result is a system of algebraic expressions over R with the structure of a field.

After this statement has been verified we have a list of theorems ready made concerning operations on algebraic expressions. They are the theorems derived from the field axioms, with "real number" replaced by "algebraic expression." Such theorems are the basis for all formal manipulations of algebraic expressions. From them we obtain such results as:

$$x^2 - 2ax + a^2 - b^2 = (x - a + b)(x - a - b)$$

$$\frac{4 - x^2}{x - 2} = -(x + 2)$$

$$\frac{1}{|x + a|} = \sqrt{\left(\frac{1}{x + a}\right)^2}$$

$$\frac{x - y}{4y^2 - x^2} - \frac{3}{x + 2y} = \frac{-4x + 7y}{x^2 - 4y^2}, \quad \text{etc.}$$

This is the bread and butter of elementary algebra; it involves the skills that every beginning student should acquire. There is a basis for these manipulations -- they constitute the art of "symbol pushing" with a purpose and for a reason -- which should take them out of the realm of mechanical busy work for a student. That is, "symbol pushing" is really concerned with the structure of the field of algebraic expressions.

Just as the system of real numbers has interesting subsystems, so has the system of algebraic expressions over a set S.

Definition: The sub-system of algebraic expressions over S obtained by applying only operations (1), (2), (4) of the definition is called the system of rational expressions over S.

For example*,

$$\frac{3x^3a - xy}{\sqrt{2} - x} \quad \text{and} \quad 5\,xy - a^2b + \pi$$

are rational expressions over R, whereas

$$\frac{3ab}{\sqrt{x - a}}$$

is an algebraic expression over I which is not rational because it involves operation (3) of extracting a root.

---

[*] Notice that $\sqrt{2}$ is an element of R and is not thought of as requiring operation (3) of the definition.

Definition: The sub-system of rational expressions over S obtained by using only the subset of connectives +, -, × is called the system of *polynomials* over S. A *monomial* is a polynomial obtained by using only the connective ×.

For example,

$3x^2a - 4yb$ is a polynomial over I,

$\frac{4x}{5} + \frac{3a^2bc}{4}$ is a polynomial[*] over F, and

$\sqrt{2}\,x^2 - 5$ is a polynomial over R.

Of special interest are polynomials in one indeterminate over R. These are of the form

$$a_0 + a_1x + a_2x^2 + \ldots + a_nx^n,$$

where $a_0, a_1, \ldots, a_n$ represent elements in R, $a_n \neq 0$, and n is a positive integer. A good deal of attention is given to such polynomials in a first course. A student learns to add and multiply polynomials, making use of associative, commutative and distributive properties. When he learns to factor polynomials (that is, into products of polynomials) he finds that he must be careful to specify the set over which the factorization takes place. For example, the polynomial over F,

$$x^2 - \frac{1}{4},$$

is not factorable into polynomials over I, but is factorable over F; whereas the polynomial over I,

$$x^2 - 2,$$

---

[*] The rational numbers $\frac{4}{5}$ and $\frac{3}{4}$ are *multiplied* by other expressions and no division is involved.

is factorable into polynomials over R but not over F or I. To illustrate the necessity of specifying the set over which factorization takes place, let us find the factors of the polynomial

$$x^4 - \frac{1}{9}.$$

(a) It is not factorable over I.

(b) Its factors over F are $(x^2 - \frac{1}{3})(x^2 + \frac{1}{3})$.

(c) Its factors over R are $(x - \sqrt{\frac{1}{3}})(x + \sqrt{\frac{1}{3}})(x^2 + \frac{1}{3})$.

(d) Its factors over the set of complex numbers are

$(x - \sqrt{\frac{1}{3}})(x + \sqrt{\frac{1}{3}})(x - i\sqrt{\frac{1}{3}})(x + i\sqrt{\frac{1}{3}})$.

As a student works with polynomials he notices their similarity to integers. Like the integers, the set of polynomials is closed under addition, subtraction and multiplication, but not under division. And also like integers, they have the property of unique prime factorization over a given set. This suggests why some writers call polynomials integral expressions.

If A and B are polynomials over S, then their quotient $\frac{A}{B}$ is certainly a rational expression over S. Conversely, it can also be shown that every rational expression over S can be represented as the quotient of two polynomials over S. The analogy with rational numbers suggests the adjective "rational" for such expressions.

Exercises

1. Identify each of the following expressions over S as polynomial, rational, algebraic, and specify the set S:

(a) $\sqrt{5x} - y + a - 2$ (d) $\frac{3}{x} + \sqrt{\frac{y}{2}}$

(b) $\frac{\pi}{3}y - \frac{9x}{\frac{x-2}{y-2}}$ (e) $\frac{1}{3} - \sqrt{\frac{1}{3}}\,xy^2$

(c) $\frac{|3 - x|}{4x}$ (f) $\frac{xy}{x - a} - \frac{by}{y - b}$

2. Represent each of the rational expressions in Problem 1 as the quotient of two polynomials.

3. If A is an algebraic expression, explain why $A - A = 0$ and $\frac{A}{A} = 1$.

4. Use the field properties of algebraic expressions to prove:

(a) $\frac{4 - x^2}{x - 2} = -(x + 2)$

(b) $\frac{x - y}{4y^2 - x^2} - \frac{3}{x + 2y} = \frac{-4x + 7y}{x^2 - 4y^2}$

(c) $\frac{x^3 + 2}{x + 1} = x^2 - x + 1 + \frac{1}{x + 1}$

5. Factor each of the following polynomials, if possible, over I, over F, over R:

(a) $x^5 - 7x^4 + 12x^3$ (c) $y^2 - 2 + a^2 - 2ay$

(b) $x^3 - \frac{5}{2}x^2 + x$ (d) $a^4 + 4$

---

3. Open Sentences. Why do we bother to construct the system of algebraic expressions over R? What good are they? Here we observe the reason for demanding that algebraic expressions satisfy the field axioms. For if we substitute real numbers for the indeterminates of an algebraic expression then the expression represents a real number (barring division by zero and roots of negatives).

Hence, with certain restrictions, we may shift back and forth between the indeterminate meaning of "variable" and the quantified meaning with certainty that no confusion will result -- under both meanings the structure of the resulting system is that of a field. The power of our algebraic manipulations comes from the fact that we may indiscriminately put on and take off quantification of the variables, in the one case talking about real numbers and in the other about indeterminates, respectively. Thus, what might appear to be laxity in an argument is really our assurance that operationally the next line of the argument will be justified whether the variables are quantified or not. This is the justification of "symbol pushing"; the power of symbol pushing comes from this freedom from specification.

The compatability of the meanings of "variable" leads to a successful marriage of their uses in algebra. Consider first the problem of solutions of sentences.

If an algebraic expression has its variables quantified with respect to the elements of a particular set T of real numbers, we say that the expression is an open phrase whose variables have domain T. As a special case we consider any element of T to be an open phrase. From open phrases we construct open sentences.

Definition: If A and B are open phrases, then "$A = B$", "$A \neq B$", "$A < B$" are open sentences. If p, q are open sentences, then "p and q", "p or q", "if p, then q" are open sentences.

Since open phrases are symbols for numbers in a given set, an open sentence is a statement concerning equality or order of numbers in this set. For example,

(1) $(x^2 - 2)(2x + 1) = 0$, x in F, and

(2) $2y < 5 - x$, x and y in N,

are open sentences, the first in one variable and the second in two. Notice particularly the quantifications of the variables; (1) is a statement about equality of rational numbers, and (2) concerns the order of natural numbers.

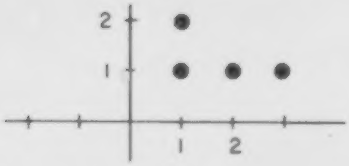
For given values of the variables an open sentence becomes a statement about numbers which is either true or false but not both. If $x = -2$, then (1) is a false statement; if $x = -\frac{1}{2}$ then it is a true statement. If $x = 2$ and $y = 1$, then (2) is a true statement, whereas if $x = 1$ and $y = 2$, it is a false statement.

A number in F which x may represent to make (1) a true statement is called a solution of (1); the set of all solutions of (1) is called the truth set (or solution set) of (1). (A discussion of truth sets of sentences in one variable was given in Chapter 2.) Thus, the truth set of (1) is $\{-\frac{1}{2}\}$. Notice that if x had instead the domain R, then the truth set of (1) would be $\{-\frac{1}{2}, \sqrt{2}, -\sqrt{2}\}$.

Before we can define the truth set of a sentence in two variables we must agree upon an order of the variables and we must construct a set of ordered pairs of numbers. In sentence (2), for example, x is given as the first variable and y the second. Furthermore, since x and y have domain N, we must construct the set of all possible ordered pairs of numbers in N: $\{(1,1), (1,2), (1,3), \ldots, (2,1), (2,2), (2,3), \ldots, (3,1), (3,2), (3,3), \ldots\}$. We call this set the cartesian product $N \times N$. Now an element in N × N (an ordered pair of natural numbers) is a solution of (2) if when x represents the first number of the ordered pair and y the second, the resulting statement is true. The set of all solutions of (2) is its truth set. Thus, the truth set of (2) is the set of two elements {(1,1), (2,1)}.

Definition: Given a sentence in two ordered variables x and y with x in S and y in T, form the set S × T of all ordered pairs, the first element of each pair being an element of S and the second an element of T. Then an element of S × T is a solution of the sentence if, when x represents the first element and y the second, the sentence is a true statement. The set of all solutions is the truth set of the sentence.

In Chapter 2 we found that graphs of sets are useful in finding and expressing truth sets of sentences in one variable. We can extend these techniques to obtain graphs of sets of ordered pairs of numbers, where the first variable has a value corresponding to a point on a horizontal number line and the second variable to a point on a vertical number line, forming the well-known cartesian coordinate system on the plane. Then each ordered pair of numbers corresponds to a unique point of the plane. For example, the set {(1,1), (1,2), (2,1), (3,1)} has the graph:



If the variables have domain R we can also say that each point of the plane corresponds to a unique ordered pair of real numbers. This one-to-one correspondence between the set $R \times R$ and the set of all points of the coordinate plane is guaranteed by the completeness of R, and this fact is the basis for analytic geometry of two dimensions.

It is an easy transition from two to three ordered variables and from two to three dimensions. The reader is invited to describe the set $R \times R \times R$ of all ordered triples of real numbers and to define the truth set of a sentence in three ordered variables, giving an appropriate description of the graph of such a truth set.

Before beginning a formal discussion of solutions of sentences we should point out the role of quantification of the variables. Consider the sentence

$$2y \leq 5 - x \text{ and } 2y \leq 2x + 5 \text{ and } y \geq 0,$$

where x and y are quantified as follows:

(a) x in N and y in I,

(b) x in I and y in N,

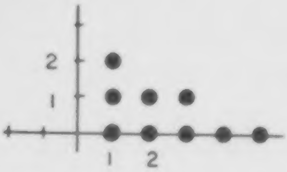
(c) x in I and y in R,

(d) x in R and y in I,

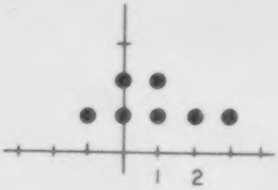
(e) x in R and y in R.

In each case the truth set of the sentence depends on the quantification. The graphs of the sentences are:

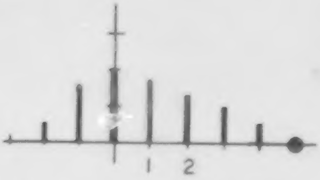
(a)



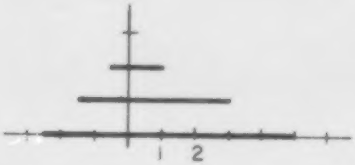
(b)



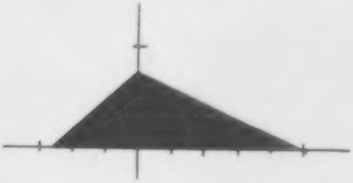
(c)



(d)



(e)



An example of the interplay of the algebras of real numbers and algebraic expressions is found in solving sentences in one variable. By solving a sentence we mean the process of determining its truth set. By a solution of a sentence we mean an element of its truth set. It requires little insight for a student to see that 2 is a solution of $x + 2 = 2x$, x in R. That 2 is the only solution can be shown by a simple argument: for any x greater than 2, $x + 2 < 2x$; for any x less than 2, $x + 2 > 2x$. Hence, the truth set is $\{2\}$. But such arguments become more difficult as the sentences become more complex. Having found solutions by trial and error, how can one be sure he has found all the solutions?

Consider, for example, the sentence

(3) $-3x^2 + 4x + 7x^2 - 5 - 3x = -x^2 + 5x + 4x^2 - 7$, x in R.

If we regard each member of the equation as an algebraic expression, we may write formally

$$-3x^2 + 4x + 7x^2 - 5 - 3x = 4x^2 + x - 5,$$
$$-x^2 + 5x + 4x^2 - 7 = 3x^2 + 5x - 7.$$

Then (3) is true if and only if

(4) $$4x^2 + x - 5 = 3x^2 + 5x - 7, \quad x \text{ in } R,$$

is true. We are assured of this because we used certain field properties in simplifying the algebraic expressions and these same properties hold for all real numbers. Thus the truth set of (4) is the same as that of (3). We say that two sentences are equivalent if they have the same truth set. Now we shift back to quantified variables and remark that by adding $(-3x^2 - 5x + 7)$, x in R, to both members of (4), the resulting sentence

(5) $$x^2 - 4x + 2 = 0, \quad x \text{ in } R,$$

is equivalent to (4). (Why is this true?) Then back to unquantified variables to factor the left member:

$$\begin{aligned} x^2 - 4x + 2 &= x^2 - 4x + 4 - 2 \\ &= (x - 2)^2 - 2 \\ &= (x - 2 - \sqrt{2})(x - 2 + \sqrt{2}), \end{aligned}$$

giving the equivalent sentence

(6) $$(x - 2 - \sqrt{2})(x - 2 + \sqrt{2}) = 0, \quad x \text{ in } R.$$

Next we use the theorem concerning real numbers (quantified variables): "a and b in R, $ab = 0 \Longleftrightarrow a = 0$ or $b = 0$", to write the equivalent sentence

(7) $$x - 2 - \sqrt{2} = 0 \text{ or } x - 2 + \sqrt{2} = 0, \quad x \text{ in } R.$$

It is an easy step to the final equivalent sentence

(8) $$x = 2 + \sqrt{2} \text{ or } x = 2 - \sqrt{2}, \quad x \text{ in } R,$$

whose truth set is, of course $\{2 + \sqrt{2}, 2 - \sqrt{2}\}$.

This solution took us through six sentences, each equivalent to the others, until we arrived at one whose truth set is obvious. In some steps we dropped the quantification and performed formal operations on algebraic expressions. At others we picked up the quantification and applied theorems concerning real numbers. Always we were assured of an equivalent sentence because the field properties we used hold true for all real numbers as well as all algebraic expressions.

The preceding solution is not intended as a model to be followed. It is a typical example in which we spell out the shifting between the two meanings of "variable". It should be noted that the factorization over R leading to (6) is accomplished by the familiar "completion of the square".

Not all operations on algebraic expressions lead to equivalent sentences. Notice that the sentence

$$\frac{x^2 - 4}{x - 2} = 4, \quad x \text{ in } R, \tag{9}$$

has a null truth set (no value of x in R makes this sentence true). But if we drop the quantifier and write

$$\frac{x^2 - 4}{x - 2} = x + 2,$$

the resulting sentence

$$x + 2 = 4, \quad x \text{ in } R, \tag{10}$$

has truth set $\{2\}$. In this case, (9) and (10) are not equivalent sentences because the formal operation, when quantified, becomes

$$\frac{x^2 - 4}{x - 2} = x + 2 \quad \text{and} \quad x \neq 2;$$

that is, we must prohibit division by 0. Now the sentence

(11) $x + 2 = 4$ and $x \neq 2$, x in R,

is equivalent to (9); it also has a null truth set.

Thus, formal operations on algebraic expressions lead to equivalent sentences if the results of the operations are then properly quantified. We assume that when theorems on real numbers are used to obtain new sentences, the quantifiers will be carefully retained.

Of course, some theorems lead to new sentences whose truth sets include those of the original sentence as proper subsets. This sometimes happens when we use the theorem: "a and b in R, and $a = b \Longrightarrow a^2 = b^2$". Note that the converse of this theorem is not true. For example, the sentence

(12) $\sqrt{2 - x} = x$, x in R,

has its truth set included in the truth set of

(13) $2 - x = x^2$, x in R.

The truth set of (13) is $\{-2,1\}$, but -2 is not a solution of (12). When we apply a theorem that does not guarantee an equivalent sentence, that is, whose converse is not true; we must check individually each member of the resulting truth set in the original sentence. On the other hand, some theorems may result in new sentences with smaller truth sets than the original sentence (such as "a and b in R, $a = b \Longrightarrow \frac{a}{c} = \frac{b}{c}$", where c involves a variable). It is best to avoid this situation if, for some x, $c = 0$.

Of primary importance to a student is his understanding of the role of equivalent sentences in solving a sentence and the types of operations and theorems resulting in equivalent sentences. Equivalence is a two-way affair. It means that every solution of the first sentence is a solution of the second, and every solution of the second is a solution of the first. If he sees how operations on algebraic expressions aid him in this procedure, he will not be tempted to treat such operations lightly.

Exercises

1. Solve each of the following sentences:

   (a) $(x + 3)(2x - 1)(x^2 - 3) = 0$, x in I.

   (b) $(x + 3)(2x - 1)(x^2 - 3) = 0$, x in F.

   (c) $(x + 3)(2x - 1)(x^2 - 3) = 0$, x in R.

   (d) $3x - 4 \leq x$, x in N.

   (e) $3y < 6 - x$, x in N, y in N.

   (f) $3y < 6 - x$ and $y \leq x$, (x,y) in $N \times N$.

   (g) $\sqrt{x - 8} = 4 + \sqrt{x}$, x in R.

   (h) $x^2 \geq 4(x - 1)$, x in F.

   (i) $\frac{|x - 1|}{x - 1} = 1$ or $x > 2$, x in R.

2. Draw the graph of:

   (a) $x^2 - 2xy = 0$, (x,y) in $R \times R$.

   (b) $3x - 2 = 0$ and $y = 4x - 1$, (x,y) in $F \times F$.

   (c) $3y < 4x + 6$ and $y < 2$ and $2y > x$, (x,y) in $I \times R$.

   (d) $3y < 4x + 6$ and $y < 2$ and $2y > x$, (x,y) in $R \times R$.

   (e) $|x| + |y| \leq 4$, (x,y) in $I \times I$.

   (f) $x^2 + y^2 < 4$ and $x > y$, (x,y) in $R \times R$.

3. Solve (by constructing a sequence of equivalent sentences):

(a) $4 + 3x^3 - 2x + 5x^2 - x^3 = 3x + 2x^3 + 2 + 2x$, x in R.

(b) $(x + 1)(x^2 - 1) = 3(x^2 - 1)$, x in R.

(c) $\frac{1}{x} + \frac{1}{1 - x} + \frac{1}{1 + x} = 0$, x in R.

(d) $(\frac{x}{x + 1})(x^2 - 1) = 0$, x in R.

(e) $\frac{x - 2}{x^2 - x - 6} = \frac{4}{x^2 - 4} + \frac{3}{2(x + 2)}$, x in R.

---

4. Functions. Running through all our discussions of operations, correspondences, algebraic expressions, and open sentences there is a common idea which was hinted at many times but never stated explicitly. This is the concept of a function. There are sharp differences of opinion on the question of introducing functions at the beginning of a first course versus ending a first course with functions. Some writers believe that all terminology of operations, correspondences, etc., should be abandoned and these ideas unified from the very beginning in terms of functions. The writers of SMSG-F, on the other hand, decided to lay the groundwork for functions and then culminate and summarize the course by showing how functions can unify the preceding ideas. The question has by no means been settled, and the reader is invited, after reading this chapter, to enter the argument, either pro or con.

If we review the major ideas of algebra, we recall such statements as

(1) Operations: For each pair of numbers a and b in R there is a unique number $a+b$ in R.

This operation assigns to *each* *pair* of elements in R exactly *one* element in R.

Each element a in N has a unique reciprocal $\frac{1}{a}$ in R.

This operation assigns to *each* element of N exactly *one* element $\frac{1}{a}$ in R.

(2) Correspondences: There is a one-to-one correspondence between the set of even natural numbers and N.

This correspondence assigns to *each* element n in N exactly *one* element $2n$ in N and assigns to *each* element e in E (even natural numbers) exactly *one* element n in N.

(3) Algebraic expressions: $3x^2 + x - 2y^2$, (x,y) in R×R. This quantified algebraic expression assigns to *each* element (x,y) in R×R exactly *one* element $(3x^2 + x - 2y^2)$ in R.

(4) Variables: Let x be the number of feet in the length of a rectangle.

The variable x assigns to *each* rectangle in the set of all rectangles exactly *one* number (of feet in its length) in R.

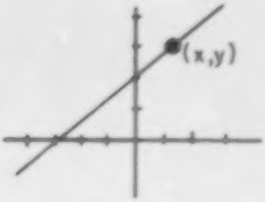
(5) Open sentences: $y = \sqrt{x}$, x in N and y in R.

This sentence assigns to *each* element x in N exactly *one* element y in R (for which the sentence is true).

(6) Sets of ordered pairs: $\{(0,2),(1,2),(3,\pi),(4,\sqrt{2}),(5,\pi)\}$. This set of ordered pairs assigns to *each* element in $\{0,1,2,3,4,5\}$ exactly *one* element in R.

(7) Graphs of sentences:

This graph assigns to *each* element x (abscissa) in R exactly *one* element y (ordinate) in R.



It is evident that a common concept runs through the above examples. In each of (1) to (7) some rule or operation or association or correspondence assigns to *each* element in a given set a *unique* element in R, resulting in a pairing off of elements from the two sets in such a way that *no two distinct elements of the second set are assigned to the same element of the first set*. To be sure, there are correspondences which pair off elements of non-numerical sets, such as the correspondence of each human being with a color (of his hair). In fact, wherever "of" or a possessive form of a verb is used there is a correspondence between elements of two sets. But in this study we shall restrict our attention to the types of correspondences given by the

Definition: Given a set of numbers and a rule which assigns to each number in this set exactly one number in R, the resulting association of numbers is called a *function*. The given set is called the *domain* of definition of the function, and the set of assigned numbers in R is called the *range* of the function.

A function is usually designated by a letter, such as f; if f assigns to each element in S exactly one element in R, we indicate this fact in various ways:

$$f\colon x \longrightarrow y, \quad x \xrightarrow{f} y, \quad f(x) = y, \quad \big(x, f(x)\big); \quad x \text{ in } S,\ y \text{ in } R.$$

The third of these notations is most commonly used in a first course. It is read "f of x is equal to y"; that is, the number assigned by the f-function to x is y. Notice that f(x) is not "f times x", but rather f(x) is a number. The fourth notation indicates that each x is paired off with the unique number f(x) assigned to x by the f-function.

A common misconception among students is that functions cannot be defined, in fact do not exist, unless there is a formula (algebraic expression) involved in the definition. We must convince him that a function is a concept, an idea, and not a formula. There are many ways of representing a function. For example, the function described above in example (5) can be represented variously by:

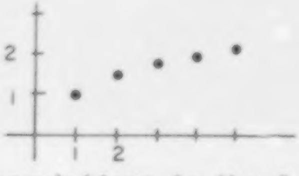
A set of order pairs: $\{(1,1),(2,\sqrt{2}),(3,\sqrt{3}),(4,2),\ldots\}$

A verbal statement: To each x in N assign the number $\sqrt{x}$ in R.

An equation: $y = \sqrt{x}$, x in N, y in R.

A formula: $f\colon x \longrightarrow \sqrt{x}$, x in N.

A graph:



No one of these representations is the function, but each describes the function. The point is that a function does not depend for its definition on its representation but only on its domain of definition and its rule of assignment. In general, two functions are *equal* if their domains are the same set and their rules of assignment are the same, regardless of the manners in which they are represented. For example, consider the two functions:

$$f:\ a \longrightarrow 2a + 1, \quad a \text{ in } I$$

$$g:\ x \longrightarrow 2x + 1, \quad x \text{ in } R$$

These are different functions because they have different domains, even though their rules of assignment are the same.

Frequently the rule of assignment is given for a function without mention of a specific domain of definition. In such a case the domain is understood to be the largest set of real numbers to which the rule can be applied sensibly. For example, if a function is defined as $f:\ x \longrightarrow \sqrt{x + 2}$, then unless otherwise stated, the domain is understood to be the set of all real numbers greater than or equal to -2.

Not all correspondences between sets of numbers define functions. This is another point of confusion for students. For example, the equation

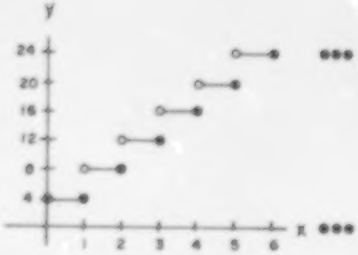
$$y^2 = x, \ x \text{ and } y \text{ in } R,$$

does *not* define a function $f:\ x \longrightarrow y$ because to each element x in R this equation assigns *two* elements y and $-y$ in R. Of course, we may write

$$y^2 = x \Longleftrightarrow y = x \quad \text{or} \quad y = -x$$

and regard the equation as defining two functions. This is precisely how we would handle this equation in certain situations in the calculus. On the other hand, the equation $y^2 = x$, x and y in R, *does* define a function g, where $g(y) = y^2$.

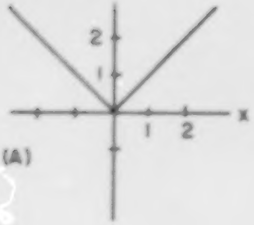
Some functions may not be represented by formulas, but this does not disqualify them as functions. For example, the first class postage regulations define a function: to each real number x (in ounces) in the set $\{0 < x \leq 320\}$ there is assigned a natural number y (in cents) according to the graph



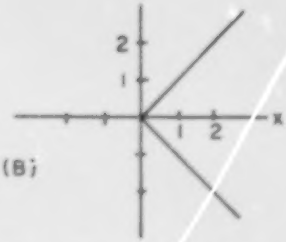
We could also describe this function verbally or represent it as a table of pairs of numbers, but we cannot find a single algebraic expression which represents y for a given x. Nevertheless, there is a function defined.

From our new point of view we can say that an *algebraic expression, when its variables are quantified, defines a function*.

The *graph* of a function f is the graph of the truth set of the sentence $y = f(x)$, with x in the domain of f. Thus, if a is in the domain of f, then $(a, f(a))$ is a point on the graph of f. From the definition of a function we see that there cannot be two points on the graph of f with the same abscissa and distinct ordinates. This is the same as saying that if a vertical line is drawn through the graph of f it will intersect the graph in exactly one point. Thus, the graph in figure (A) describes a function, whereas the graph in figure (B) does not.



(A)



(B)

For a student the graphical representation of a function is probably more informative than any other. For instance, figure (A) gives the graph of the absolute value function defined by the equation

$$f(x) = |x|.$$

From the graph it is easy to see that another representation is given by

$$f(x)= \begin{cases} x, \text{ if } x \geq 0, \\ -x, \text{ if } x < 0, \end{cases}$$

which is a statement of the definition of $|x|$.

The study of linear and quadratic functions is aided by graphs, and the subtleties of domains of definition are often cleared up by graphical representation.

A final word to teachers. When students are introduced to functions, the introduction must be clear and precise. It would be better to omit all mention of functions rather than present a vague meaning of them. But if a student really understands what a function is he can begin to see the unity and coherence of the variety of topics he studied in algebra.

Exercises

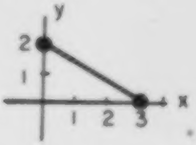
1. Each of the following is a representation of a function; give three other representations and describe its domain and range.

(a) To each positive integer n there is assigned its remainder after dividing n by 5.

(b) $\{(1,5),(2,8),(3,11),(4,14),(5,17),\ldots\}$

(c) $f(x) = |x + 2|, \quad x \text{ in } \{-3,-2,1,2\}$.

(d)



2. Determine the domain of each of the functions defined as follows:

(a) $f(x) = \frac{1}{x} - x^2$

(c) $h(x) = \sqrt{\frac{x}{x - 1}}$

(b) $g(x) = \sqrt{x^2 - 4}$

(d) $k(x) = \sqrt{|x + 1|}$

3. How are the functions in each of the following pairs related?

(a) $f(x) = x - 2, \quad F(t) = \frac{t^2 - 4}{t + 2}$

(b) $g(x) = x^2 - 1, \quad G(t) = \frac{t^4 - 1}{t^2 + 1}$

(c) $h(x) = \sqrt{(x - 1)^2}, \quad H(t) = |t - 1|$

4. Consider the function f defined by the rule

$$f(x) = \begin{cases} -1, \text{ if } -1 \leq x < 0 \\ x, \text{ if } 0 < x \leq 2 \end{cases}$$

(a) What numbers are represented by $f(-\frac{1}{2})$, $f(\sqrt{5})$, $f(\frac{3}{2})$?

(b) What is the domain of f?

(c) What is the range of f?

(d) What is the truth set of the equation $f(x) = x$?

(e) Draw the graph of the truth set of the sentence $f(x) < 1$.

5. Given the function g defined by

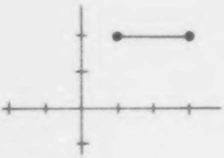
$$g(x) = x^2 - 1, \quad x \text{ in } R.$$

If t is in R, what numbers are represented by $g(-t)$, $-g(t)$, $2\,g(t)$, $g(2t)$, $g(t-1)$, $g(t) - 1$, $g\big(g(t)\big)$, $g\big(g(\frac{1}{t})\big)$, $g(\frac{1}{g(t)})$?

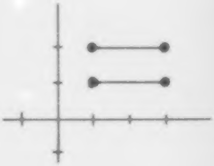
6. Draw the graph of a function f which satisfies the conditions $f(-1) = -2$, $f(0) = f(1) = 0$, $f(2) = 2$, $f(x) < 0$ for $-1 < x < 0$, and $f(x) > 0$ for $0 < x < 1$ and for $1 < x < 2$. Is there only one function satisfying these conditions?

7. Which of the following graphs define functions?

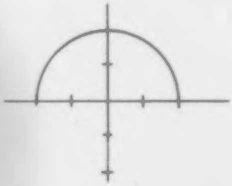
(a)



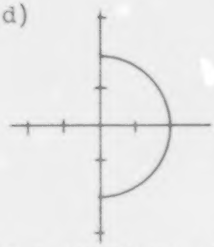
(b)



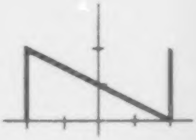
(c)



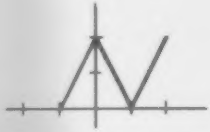
(d)



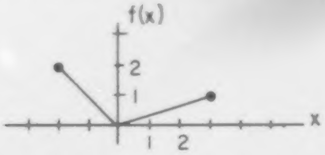
(e)



(f)



8. If a function f is defined by the following graph



draw the graphs of the following functions:

(a) g, where $g(x) = -f(x)$, $-2 < x < 3$

(b) h, where $h(x) = f(-x)$, $-3 < x < 2$

(c) k, where $k(x) = f(x) + 2$, $-2 < x < 3$

(d) t, where $t(x) = f(x + 2)$, $-4 < x < 1$.

Appendix A

INFINITE DECIMALS

1. *Decimal Representations of Real Numbers*. It was established in Chapter 4 that integers and rational numbers may be represented as decimals, the former always terminating and the latter sometimes non-terminating (*infinite*). It was also shown that if a rational number has an infinite decimal representation, it is *periodic* in the sense that its digits repeat in a regular fashion. This leaves unanswered the question: Is every periodic decimal the representation of a rational number? After we answer this in the affirmative, there is raised the new question: What numbers, if any, are represented by non-periodic infinite decimals?

Recall that a positive number x is *represented* by an infinite decimal if x satisfies *every* one of the inequalities in the infinite set of inequalities:

$$a_k \leq x \leq b_k, \quad a_k < b_k, \quad b_k - a_k = \frac{1}{10^k}, \quad k = 1,2,3, \ldots ,$$

where

$$a_k = c_n 10^n + \ldots + c_o + \frac{d_1}{10} + \ldots + \frac{d_k}{10^k}, \quad b_k = c_n 10^n + \ldots + c_o + \frac{d_1}{10} + \ldots + \frac{d_k + 1}{10^k} ,$$

n is in N, and each c_i and d_i is some integer in $\{0,1,\ldots,9\}$. Then we write $x = c_n \ldots c_o . d_1 \ldots d_k \ldots$

We shall show that *all real numbers have decimal representations* and *all decimals represent real numbers*. Then, if a real number is *not* rational its decimal representation is *not* periodic. Conversely, since every rational number is represented by a terminating or a periodic decimal, if a decimal is neither terminating nor periodic it must represent a real number which is not rational. Here we have a clear distinction between rational and irrational numbers -- a distinction of *periodicity* of their decimal representations.

The above remarks need to be verified. First, let us show that all decimals represent real numbers. Of course, if a decimal terminates, then by definition it is the *rational* sum of a finite set of rational numbers each of the form

$$c_i 10^i \quad \text{or} \quad \frac{d_i}{10^i}, \quad 0 \leq c_i \leq 9, \quad 0 \leq d_i \leq 9.$$

Thus it is sufficient to show that all infinite decimals represent real numbers. For example, consider the infinite decimal .232332333233332... and the corresponding infinite set of terminating decimals

$$D = \{.2, .23, .232, .2323, .23233, .232332, .2323323, \ldots\}$$

where each terminating decimal in D contains one more digit than the preceding. Certainly D is a non-empty set of rational numbers which has an upper bound, say 1. Hence, by axiom C, D has a *unique least upper bound which is a real number*. We shall show that this unique lub of D is the real number represented by the infinite decimal .2323323332... . Consider the infinite set of inequalities:

$$\begin{aligned} a_1 &= .2 < .3 = b_1 \\ a_2 &= .23 < .24 = b_2 \\ a_3 &= .232 < .233 = b_3 \\ &\vdots \qquad \vdots \qquad \vdots \end{aligned}$$

Since $a_n < b_n$, $a_n \leqq a_{n+1}$, $b_{n+1} \leq b_n$, and $b_n - a_n = \frac{1}{10^n}$ for all n in N, there is exactly one real number which satisfies every inequality, and this number is the lub of D. (See Problem 5 on p. 5.19.) In general, corresponding to each infinite decimal $.d_1d_2\ldots d_k\ldots$ there is an infinite set of terminating decimals

$$D = \{.d_1, .d_1d_2, .d_1d_2d_3, \ldots, .d_1d_2\ldots d_k, \ldots\},$$

$0 \leq d_i \leq 9$, which is bounded above and has a lub which is the unique real number represented by the infinite decimal.

Conversely, every real number can be represented by a decimal. This has been shown for rationals; it remains to be shown for irrationals. We first recall that any irrational number may be approximated as closely as desired by a rational number (see Problem 4 on page 5.19); for our purposes let us state this fact in the following form: Given any irrational number y and any positive rational number e, no matter how small, there is a rational number x such that

$$x < y < x + e.$$

Now it is possible to generate a set of successive rational approximations a_k to y corresponding to the successive values of e: 1, .1, .01, .001, ... Hence, we have a set of inequalities

$$\begin{array}{l} a_0 < y < a_0 + 1 = b_0 \\ a_1 < y < a_1 + \frac{1}{10} = b_1 \\ a_2 < y < a_2 + \frac{1}{10^2} = b_2 \\ \vdots \quad \vdots \qquad\quad \vdots \end{array}$$

and for every k in N, $a_k < y < a_k + \frac{1}{10^k} = b_k$, where $a_k < b_k$, $a_k \leq a_{k+1}$, $b_{k+1} \leq b_k$, $b_k - a_k = \frac{1}{10^k}$. This set of rational approximations a_k will be bounded above and will have y as its lub. The corresponding infinite decimal will represent y.

For example, let us find the infinite decimal representation of the irrational number $\sqrt{2}$. Corresponding to e = 1,

$$1 < \sqrt{2} < 1 + 1, \text{ since } 1 < 2 < 4.$$

Corresponding to e = .1,

$$1.4 < \sqrt{2} < 1.4 + .1, \text{ since } (1.4)^2 < 2 < (1.5)^2.$$

e = .01: $\quad 1.41 < \sqrt{2} < 1.41 + .01$, since $(1.41)^2 < 2 < (1.42)^2$.

e = .001: $\quad 1.414 < \sqrt{2} < 1.414 + .001$, since $(1.414)^2 < 2 < (1.415)^2$.

etc.

The lub of the set $\{1, 1.4, 1.41, 1.414, \ldots\}$ is the irrational number $\sqrt{2}$, and is represented by the infinite decimal 1.414... .

It is not surprising that the completeness axiom provides the answer to the problem of decimal representation of real numbers. After all, it is this axiom which completes the characterization of R. However, it must be noted that we have stated an *existence* theorem for decimal representation of an irrational number. It tells us there *is* a decimal, but it does not spell out a method for finding the particular digits of the decimal. For square roots there are algorithms of various kinds which exhibit the digits of the representation, but for such irrational numbers as π, e, log 2, etc., we must develop special methods for each number.

Now we are assured that any infinite decimal, say

$$.3212121\ldots = .3\dot{2}\dot{1},$$

represents a real number. It remains to show that *it is a rational number if its representation is periodic*. To illustrate the technique we shall use, let us consider the periodic decimal $.3\dot{2}\dot{1}$. Let .212121... represent a real number N so that

$$.3212121\ldots = \frac{3}{10} + \frac{N}{10} .$$

Now

$$\begin{aligned} N = .2121\ldots &\Longrightarrow 100N = 21.2121\ldots \\ &\Longrightarrow 100N = 21 + N \\ &\Longrightarrow 99N = 21 \\ &\Longrightarrow N = \frac{7}{33} . \end{aligned}$$

Hence, $.3212121\ldots = \frac{3}{10} + \frac{7}{330} = \frac{53}{165}$, which is a rational number. Long division will verify the periodicity.

It is interesting to study the decimal $.999\ldots9\ldots = .\dot{9}$. If $N = .999\ldots$, then $10N = 9.999\ldots = 9 + N$. Hence, $9N = 9$ and $N = 1$. Thus we see that the rational number 1 can be represented by either 1.000... or by .999... . This choice of two periodic decimal representations is possible for every terminating decimal:

$$.325 = .325000\ldots = .34999\ldots = .34\dot{9}$$

$$4.728 = 4.728\dot{0} = 4.727\dot{9}.$$

The technique used above can be applied to any periodic decimal as follows: Let us assume that the repeating block of k digits first occurs after the jth digit to the right of the decimal point.

$$x = c_n c_{n-1} \cdots c_1 c_0 . d_1 d_2 \cdots d_j \underbrace{\dot{d}_{j+1} \dot{d}_{j+2} \cdots \dot{d}_{j+k}}_{\text{k digits}} .$$

Then x is the sum of a rational terminating decimal and a periodic decimal:

$$x = c_n c_{n-1} \cdots c_1 c_0 . d_1 d_2 \cdots d_j + \frac{1}{10^j}(.\dot{d}_{j+1} \dot{d}_{j+2} \cdots \dot{d}_{j+k}) .$$

Let $.\dot{d}_{j+1}\dot{d}_{j+2}\ldots\dot{d}_{j+k}$ represent a real number N. Then

$$10^k N = d_{j+1}d_{j+2}\ldots d_{j+k}.\dot{d}_{j+1}\dot{d}_{j+2}\ldots\dot{d}_{j+k}$$

$$10^k N = d_{j+1}d_{j+2}\ldots d_{j+k} + N$$

$$N = \frac{d_{j+1}d_{j+2}\ldots d_{j+k}}{10^k - 1}$$

Thus, N is a rational number, being the quotient of the integers $d_{j+1}d_{j+2}\ldots d_{j+k}$ and $10^k - 1$.

To summarize, we list the facts we have learned:

(1) Every decimal represents a real number.

(2) Every real number has a decimal representation.

(3) The infinite decimal $.d_1d_2\ldots d_k \ldots$ represents the lub of the infinite set D of rational numbers:

$$D = \{.d_1, .d_1d_2, .d_1d_2d_3, \ldots, .d_1d_2\ldots d_k, \ldots\}$$

(4) The decimal representation of a real number is periodic if and only if it is a rational number (writing terminating decimals as periodic decimals with repeating zeros).

(5) The representation of a real number is non-periodic if and only if it is an irrational number.

2. R is Not Countable. Now that we know all real numbers can be represented as infinite decimals, we are in a position to prove a statement made in Chapter 4: The set R is not countable. For simplicity, let us restrict our attention to the set Q of all real numbers x such that $0 < x < 1$. If we can prove this set is not countable, then certainly R is not countable.

Let us assume the negative, namely, that the set Q of all real numbers between 0 and 1 is countable, and obtain a contradiction.

This means we assume a one-to-one correspondence between the elements of N and Q. Now every real number in Q can be written as an infinite decimal, where we may agree to write terminating decimals as periodic decimals with repeating zeros. Then we form the correspondence:

$$\begin{array}{ccl} \underline{N} & & \quad\underline{Q} \\ 1 & \longleftrightarrow & .a_1a_2a_3\ldots\ a_k\ldots \\ 2 & \longleftrightarrow & .b_1b_2b_3\ldots\ b_k\ldots \\ 3 & \longleftrightarrow & .c_1c_2c_3\ldots\ c_k\ldots \\ \vdots & & \quad\vdots \\ n & \longleftrightarrow & .r_1r_2r_3\ldots\ r_k\ldots \\ \vdots & & \quad\vdots \end{array}$$

where all the digits are in the set $\{0,1,2,\ldots,9\}$. By assumption, every real number in Q is in this list. To show a contradiction let us construct a real number x in Q which cannot be listed. Form

$$x = .t_1t_2t_3\ldots t_k\ldots,$$

with not all its digits 9, where t_1 is a digit different from a_1, t_2 is different from b_2, t_3 is different from c_3, ..., t_n is different from r_n, ... Certainly, x is different from each real number listed because it differs in at least one digit from each of the numbers. Yet x must be in Q because it is represented by a positive decimal which is less than 1. This is a contradiction; hence, R is not countable.

Exercises

1. Find the rational number represented by each of the following periodic decimals:

(a) $.\dot{1}\dot{2}\dot{3}$ (c) $.00\dot{3}\dot{6}$ (e) $6.35\dot{0}$

(b) $4.3\dot{1}$ (d) $.\dot{1}\dot{4}\dot{2}\dot{8}\dot{5}\dot{7}$ (f) $.10\dot{9}$

2. Determine the first four elements of an infinite set of terminating decimals whose lub is

(a) $\sqrt{3}$ (c) $\sqrt[3]{2}$

(b) $\frac{1}{3}$ (d) $\sqrt{5}$

3. Find a terminating decimal

(a) between $\sqrt{5}$ and $\sqrt{6}$,

(Hint: $\sqrt{5} < t < \sqrt{6} \iff 5 < t^2 < 6$.)

(b) between $\sqrt{43}$ and $\sqrt{44}$,

(c) between $\frac{31}{33}$ and $\frac{32}{33}$.

4. Prove that the infinite decimal

$$.10100100010000100000 1\ldots$$

where the number of 0's between successive 1's increases as indicated, represents an irrational number.

5. Explain why neither 3.1416 nor $\frac{22}{7}$ represents the number π.

6. A real number may be represented by an infinite set of digits taken from any set of integers of the form $\{0,1,\ldots,(k-1)\}$. If $k = 2$, for example, we have binary representation. Then

$$10.11_{\text{two}} = 1\cdot 2 + 0\cdot 1 + \frac{1}{2} + \frac{1}{2^2} .$$

Find the rational numbers in the binary scale represented by each of the following periodic binary forms:

(a) $.\dot{1}$ (c) $.1\dot{1}\dot{0}$

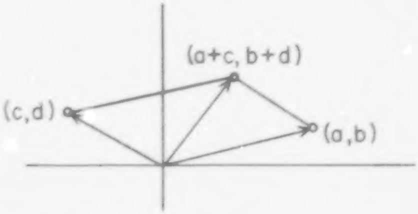
(b) $.\dot{0}\dot{1}$ (d) $1.0\dot{1}\dot{0}\dot{1}$

Appendix B

COMPLEX NUMBERS

There are some desired properties that the set R of real numbers does not have. For one, the equation $x^2 + 1 = 0$ does not have a solution in R. As long ago as the beginning of the 19th century there were attempts made to develop a number system in which such equations have solutions. In the 1840's Hamilton introduced the complex number system as follows.

Just as each point of the number line is associated with a real number, Hamilton associated each point of the plane with a complex number denoted by an ordered pair of real numbers (a,b). His initial problem was to define equality, addition, and multiplication of these "points" in such a way that the resulting system of complex numbers is a field which includes the system of real numbers as a proper subsystem. He was motivated in his definitions by the desire to have the solutions of the equation $x^2 = -1$ in this system and by the observation that complex numbers should add like vectors.



B.1

Definition: Consider the set Z of all ordered pairs (a,b), a and b in R, with "=", "+", "." defined for these elements as follows: For a, b, c, d, in R, (a,b) and (c,d) are in Z and

$$(a,b) = (c,d) \Longleftrightarrow a = c \text{ and } b = d$$

$$(a,b) + (c,d) = (a + c, b + d),$$

$$(a,b)\cdot(c,d) = (ac - bd, ad + bc).$$

The resulting system $(Z, +, \cdot)$ is called the complex number system.

The reader is invited to use the properties of operations in R to prove that the set Z is closed under the operations $+$, $\cdot$, as defined above, that these operations are commutative and associative, and that $\cdot$ is distributive through $+$. Since

$$(a,b) + (0,0) = (a,b) \quad \text{and} \quad (a,b)\cdot(1,0) = (a,b)$$

for all a, b in R, the system contains an additive identity $(0,0)$ and a multiplicative identity $(1,0)$. Also, since

$$(a,b) + (-a,-b) = (a,b) \quad \text{for all } a,b \text{ in R},$$

and

$$(a,b)\cdot\left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2}\right) = (1,0) \quad \text{for all } (a,b) \neq (0,0),$$

the system contains an additive inverse $(-a,-b)$ for each element (a,b) and a multiplicative inverse $\left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2}\right)$ for each non-zero element (a,b). We conclude that the system $(Z, +, \cdot)$ satisfies axioms F1 to F7 and is a field; all the properties proved for a field are enjoyed by the system of complex numbers.

It should be noted immediately that the particular subset R' of Z consisting of all complex numbers of the form $(a,0)$ is a very familiar set. Consider the following properties:

$$\begin{aligned}
(a,0) &= (c,0) \iff a = c && \text{for } a, c \text{ in } R,\\
(a,0) + (c,0) &= (a + c, 0) && \text{for } a, c \text{ in } R,\\
(a,0)\cdot(c,0) &= (ac,0) && \text{for } a, c \text{ in } R,\\
(a,0) + (-a,0) &= (0,0) && \text{for all } a \text{ in } R,\\
(a,0)\cdot(\tfrac{1}{a},0) &= (1,0) && \text{for all } a \neq 0 \text{ in } R.
\end{aligned}$$

We conclude that the system $(R', +, \cdot)$ is also a field. In fact, it is in every respect like the field of real numbers, for there is a one-to-one correspondence between R' and R in which the complex number $(a,0)$ corresponds to the real number a and in which the operations are preserved:

$$\begin{aligned}
(a,0) &\longleftrightarrow a\\
(a,0) + (c,0) &\longleftrightarrow a + c\\
(a,0)\cdot(c,0) &\longleftrightarrow ac, \quad \text{for all } a,c \text{ in } R.
\end{aligned}$$

Because of this operation-preserving correspondence and the fact that the systems have the same structure, we adopt the convention that R' and R are the same set, and we write a in place of $(a,0)$ whenever convenient. In this sense we have shown that the set R of real numbers is a subset of the set Z of complex numbers.

Is R a proper subset of Z? To answer this question in the affirmative let us concentrate on the element $(0,1)$ in Z. By definition,

$$(0,1)\cdot(0,1) = (0 - 1, 0 + 0) = (-1,0) = -1.$$

Hence, we have found an element in Z whose "square" is the real number -1. But we know that there is no real number whose square is -1; we conclude that $(0,1)$ cannot be identified with a real number. Thus, $R \neq Z$.

This complex number $(0,1)$ is called the imaginary unit and is denoted by i. Now we observe that

$$(a,b) = (a,0) + (0,b) \quad \text{and} \quad (0,1)\cdot(b,0) = (0,b)$$

implies

$$(a,b) = (a,0) + (0,1)\cdot(b,0) = a + ib.$$

The notation $a + ib$ for a complex number is more convenient than (a,b) because it gives us a device for remembering the definitions of addition and multiplication of complex numbers. Making use of the associative, commutative and distributive properties, we have

$$\begin{aligned}(a + ib) + (c + id) &= (a + c) + i(b + d),\\ (a + ib)\cdot(c + id) &= (ac + i^2bd) + i(ad + bc)\\ &= (ac - bd) + i(ad + bc),\end{aligned}$$

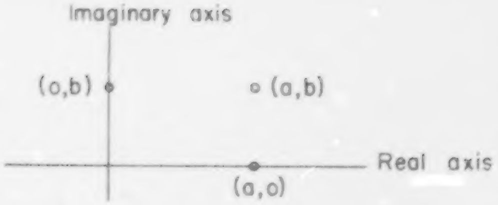
since we have shown that $i^2 = -1$. Furthermore, $i^3 = -i$, $i^4 = 1$, $i^5 = i$, $i^6 = -1$, ..., so that every number of the form

$$a_0 + a_1i + a_2i^2 + \dots + a_ni^n,$$

where a_k is in R for each $k = 1, 2, 3, \dots, n$, can be expressed in the form $a + ib$, a and b in R.

The set of complex numbers of the form (a,b), $b \neq 0$, is called the set of imaginary numbers. Thus, the set Z can be considered as the enlarged set obtained by annexing to the real numbers (complex numbers of the form (a,b), $b = 0$) the set of imaginary numbers (complex numbers of the form (a,b), $b \neq 0$).

The set Z is associated with the set of points in a plane by the simple device of referring to a pair of rectangular coordinate axes and letting each complex number (a,b) correspond to the point (a,b) in the plane, a and b in R.



Thus, the real numbers correspond to the points on the horizontal (real) axis, and the imaginary numbers correspond to the points in the plane not on the horizontal axis. It is clear that the completeness of R guarantees that every point of the real axis corresponds to a number of the form (a,0), that every point of the vertical (imaginary) axis corresponds to a number of the form (0,b), and, finally, that every point of the plane corresponds to a number of the form (a,b). All these correspondences are one-to-one.

Let us review the properties of Z:

(1) Closure. The set Z is closed under addition, subtraction, multiplication and division (excluding division by (0,0)). Of prime importance is the fact that if z is a complex number, then $\sqrt[n]{z}$ is also a complex number for any n in N. We shall not prove this here; the proof involves a different representation of complex numbers. The equation $x^2 = a$, a in Z, has a solution in Z and, more generally, every polynomial equation in one variable

$$a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 = 0,$$

a_i in Z, has a solution in Z. This remarkable theorem, called the fundamental theorem of algebra, was first proved by Gauss in 1799. Thus, no more extensions beyond the complex number system are necessary for solutions of polynomial equations. A consequence of the fundamental theorem is that every polynomial can be factored over Z. For example,

$$x^2 + 9 = (x + 3i)(x - 3i),$$

$$x^2 + 2x + 5 = (x + 1)^2 + 4 = (x + 1 + 2i)(x + 1 - 2i).$$

(2) Completeness. The system of complex numbers is complete only in the sense that there is a one-to-one correspondence between Z and the set of all points in the plane. The axiom of completeness applies only to ordered systems, and there is no way to define order for complex numbers so that the order axioms 01 to 04 hold true.

In Chapter 1 we remarked that the most important discoveries in algebra have been made by studying structures of systems without regard for the models suggested. For example, a large part of modern abstract algebra was motivated by Hamilton and Cayley in the 1840's when they looked at some known results of algebra from the point of view of structure. Their work contained one of the first illustrations of the possibility of making significant new discoveries in mathematics as a result of examining the structure of known results.

The known results at that time were the properties of real numbers. It was known that real numbers can be associated with points of a line, and that there is an ordering of the real numbers.

If the real number a is positive $(a > 0)$ or negative $(a < 0)$, then $a^2 > 0$. Thus, $a_1^2 + a_2^2 + \ldots + a_n^2 = 0$ implies that $a_1 = a_2 = \ldots = a_n = 0$, and $x^2 = b$, $b < 0$, has no solution among the real numbers. Cauchy, Gauss and others* introduced a solution i of the equation $x^2 = -1$ and, adding this "imaginary" number to the real numbers, saw that the resulting number system contains all expressions of the form

$$a + bi + ci^2 + di^3 + \ldots ,$$

all of which simplify to $r + si$, where $a,b,c,d,\ldots,r,s$ are real numbers. Moreover,

$$(a + bi) + (c + di) = (a + c) + (b + d)i$$

and

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i.$$

This was the situation when Hamilton came on the scene in 1843.

First of all he looked at complex numbers, as numbers of the form $a + bi$ were called, from the viewpoint of analytic geometry. Just as a point on the line corresponds to a single real number, so a point in the plane can be made to correspond to a single pair of real numbers (a,b). Thus, Hamilton thought of each point of the plane as a single complex number which he denoted by a number couple (a,b). His problem was: Can multiplication of points be defined in such a way that the system has the same structure as the real numbers, at least as far as addition and multiplication are concerned?

---

*See E.T. Bell, Men of Mathematics, pages 232-234.

He proceeded to define addition and multiplication of points of the plane, as was done earlier in this section, and then was able to show that the resulting system, like the real numbers, has the properties of a field, and also contains a solution of the equation $x^2 = -(1,0)$, namely, $x = (0,1)$.

He observed more. The distance from the origin to the point (a,b) is given by $\sqrt{a^2 + b^2}$; if z is the complex number (a,b), we write $|z| = \sqrt{a^2 + b^2}$ and call $|z|$ the modulus of z. Now every complex number $z = (a,b)$ and its conjugate $\overline{z} = (a,-b)$ satisfy the quadratic equation with real coefficients:

$$z^2 - 2az + a^2 + b^2 = 0.$$

Also,

$$z\,\overline{z} = |z|^2 = a^2 + b^2$$

and

$$|z_1|^2|z_2|^2 = |z_1 z_2|^2.$$

Finally, if $z_1 = (a,b)$ and $z_2 = (c,d)$, then

$$(a^2 + b^2)(c^2 + d^2) = (ac - bd)^2 + (ad + bc)^2;$$

that is, the product of two sums of two squares can be written as the sum of two squares. This result led Hamilton, Grassmann and others to ask: Can the product of two sums of n squares be written as a sum of n squares? In other words, for what values of n can we write

$$(a_1{}^2 + a_2{}^2 + \dots + a_n{}^2)(b_1{}^2 + b_2{}^2 + \dots + b_n{}^2) = A_1{}^2 + \dots + A_n{}^2,$$

where $A_1, A_2, \dots, A_n$ are certain sums and products of $a_1, a_2, \dots, a_n$ and $b_1, b_2, \dots b_n$?

Here we have the second of two important problems. The first can be generalized as follows: Call n-dimensional space the collection of all points $(x_1, x_2, \ldots, x_n)$, where each x_i is a real number, and add points according to the law

$$(a_1, a_2, \ldots, a_n) + (b_1, b_2, \ldots, b_n) = (a_1 + b_1, \ldots, a_n + b_n).$$

For what values of n is it possible to define multiplication of points

$$(a_1, a_2, \ldots, a_n)(b_1, b_2, \ldots, b_n) = (c_1, c_2, \ldots, c_n)$$

in such a way that the resulting system has the structure of a field?

Both problems were already solved for $n = 2$. Hamilton made the discovery that when $n = 4$ the first problem of defining multiplication of points in 4-space is possible and the resulting system, which he called quaternions, had all the properties of a field except for the commutative property of multiplication. In the process he also solved the second problem for $n = 4$.

Cayley in 1845 showed that both problems have a solution for $n = 8$; however, in this case neither the commutative nor the associative properties of multiplication hold.

Much effort was subsequently expended on both problems. In 1898 Hurwitz proved that the second problem has a solution only for $n = 1, 2, 4, 8$. The other problem remained open until 1940 when the Swiss mathematician Hopf used powerful new methods of algebraic topology to show that the first problem has solutions only for n a power of 2. Then in 1957, using still more refined topological methods, a solution was finally given independently by M. Kervaire and J. Milnor.

The result is that Hamilton and Cayley had found them all: the only values of n for which we can successfully define multiplication of points in n-space are $1, 2, 4, 8$.

Exercises

1. Using the definition of operations on complex numbers, prove that in the system $(Z, +, \cdot)$:
   (a) multiplication is associative,
   (b) multiplication is distributive through addition.
   (If $u = (a,b)$, $v = (c,d)$, $w = (e,f)$, then $u(v + w) = uv + uw$.)
2. Using the definition of equality of complex numbers, prove that:
   (a) $(a,b) = (c,d)$ and $(c,d) = (e,f) \Longrightarrow (a,b) = (e,f)$
   (b) $(a,b) = (c,d) \Longrightarrow (a,b) + (e,f) = (c,d) + (e,f)$
   (c) $(a,b) = (c,d) \Longrightarrow (a,b)\cdot(e,f) = (c,d)\cdot(e,f)$
3. Solve for x in Z:
   (a) $x^2 + 4 = 0$
   (b) $x^2 + x + 1 = 0$
   (c) $2x^3 - 4x^2 = 3x$
   (d) $(x - 3)(x^2 - 5)(x^2 + 9) = 0$
4. If $u = (3,-1)$, $v = (-4,2)$, $w = (0,3)$, compute
   (a) $u + v$
   (b) $\frac{v}{w}$
   (c) $u(v + w)$
   (d) $v - \frac{w}{u}$
   (e) $\bar{u}\,u$
   (f) $\overline{u\,v} - \bar{u}\,w$
   (g) $u^2 - v^2$
   (h) $\bar{u}\,\overline{\bar{v}\,w}$

5. We associated with each element $z = (a,b)$ in Z a number $|z| = \sqrt{a^2 + b^2}$ in R, called the modulus of z, which represents the distance between the points $(0,0)$ and (a,b). Show that if u and v are in Z, then:

(a) $|u| \cdot |v| = |u \cdot v|$

(b) $|u + v| \leq |u| + |v|$

(c) $|u|^2 = u \cdot \overline{u}$

(d) If we establish an order $\prec$ among elements of Z by the definition

$$u \prec v \longleftrightarrow |u| < |v|,$$

which, if any, of the order axioms are satisfied?

Appendix C

ALGEBRAIC NUMBERS

When a student visualizes the set R of real numbers he usually thinks of two subsystems, the rationals and the irrationals, which are disjoint. That is, a real number is either rational or irrational, but not both. He is usually content to let the matter rest there.

But the mathematician is forever classifying. He knows that the set of rationals is countable and the set of irrationals is not. These questions naturally come to his mind: Are there other possible classifications of the reals? Is the set F of rationals the largest countable subset of R? It turns out that his curiosity leads him to the discovery that there are other classifications of the reals, and there is a countable subset of R which contains F as a proper subset.

These results usually surprise a student. Why should he be surprised? Possibly because he has a limited experience with irrationals. When asked for an example of an irrational he will probably say, "$\sqrt{2}$" or "$\sqrt[n]{x}$, where x is an integer which is not a perfect nth power." When asked for an example of an irrational which is not obtained as a root, seldom will he respond with π. Even if he has studied logarithms and trigonometry he is not liable to give log 2 or sin 2 as an example. Somehow or other he thinks of the values of logarithmic and trigonometric functions as "different" numbers which are real but vaguely unrelated to the properties of the reals. To him the bulk of the irrationals is found among the nth roots of integers. We shall show that this is not the case.

How can we characterize numbers which are obtained as roots? By definition, $\sqrt[5]{3}$ is a solution of the polynomial equation $x^5 - 3 = 0$. $1 + \sqrt{2}$ is a solution of the polynomial equation $x^2 - 2x - 1 = 0$, as the reader may verify. These and other examples suggest a new classification of the real numbers in terms of solutions of certain polynomial equations. In the following, by "polynomials" we shall mean polynomials with integers for coefficients.

Definition: The number x is called algebraic if it is a solution of some polynomial equation

$$a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 = 0,$$

where each a_i is in I and n is in N. If x is not algebraic, it is called transcendental.

Let us here restrict our attention to the real numbers. Then a real number is either algebraic or transcendental, but not both, depending on whether or not it is a solution of some polynomial equation.

What is to be learned from such a new classification? First we notice that all rational numbers are algebraic (being solutions of $ax - b = 0$, a and b in I, $a \neq 0$) and all real numbers of the form $\sqrt[n]{a}$, a in I, $a \geq 0$, are algebraic. But some real numbers of the form $\sqrt[n]{a}$ are not rational. Thus, the set of real algebraic numbers includes F as a proper subset. But is the set of real algebraic numbers countable?

The answer is "yes". We arrive at this result as follows: First, let us accept without proof the fact that corresponding to each algebraic number A there is a unique polynomial equation of lowest degree n such that A is a solution of the equation.

For example, if A is the rational number $\frac{p}{q}$, there is a unique equation of first degree, namely $qx - p = 0$, which is satisfied by A. If $A = \sqrt[n]{a}$, there is a unique nth degree equation, $x^n - a = 0$, which is satisfied by A. In general we would follow the line of reasoning used in the following example. Consider the algebraic number $x = \frac{-13 + \sqrt{115}}{2}$.

Then $2x + 13 = \sqrt{115}$, and $4x^2 + 52x + 169 = 115$; thus

$$2x^2 + 26x + 27 = 0$$

is the polynomial equation of lowest degree, namely 2, whose solution is $\frac{-13 + \sqrt{115}}{2}$. We see that this is the lowest degree because we must square both members of the equation to obtain a polynomial equation.

Next, we define the index of the polynomial equation

$$a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 = 0$$

to be the positive integer

$$h = n + |a_n| + |a_{n-1}| + \dots + |a_1| + |a_0| .$$

Now for each positive integer h there is a finite number of polynomial equations having index h. For example, there is exactly one equation with index $h = 2$, namely, $x = 0$.

There are exactly 4 equations with index 3:

$$2x = 0, \ x + 1 = 0, \ x - 1 = 0, \ x^2 = 0.$$

There are exactly 10 equations with index 4:

$$x + 2 = 0, \quad x - 2 = 0, \quad 2x + 1 = 0, \quad 2x - 1 = 0,$$
$$3x = 0, \quad x^2 + x = 0, \quad x^2 - x = 0, \quad x^2 + 1 = 0,$$
$$2x^2 = 0, \quad x^3 = 0.$$

(Note that we are considering only real numbers and thus will discard the equation $x^2 + 1 = 0$.) How many polynomial equations have index 5?

Now we have a scheme for counting the algebraic numbers. For each successive value of $h = 2, 3, 4, 5, \ldots,$ there is a finite number of polynomial equations each with a finite number of roots which can be listed in some order. Thus, there can be established a one-to-one correspondence between N and the set of algebraic numbers. As a consequence, the set of algebraic numbers is countable and has F as a proper subset.

What are some properties of the real algebraic numbers? It can be shown that they satisfy the axioms for an ordered field but not the completeness axiom. Also, since the set of real algebraic numbers is countable, the set of real transcendental numbers is not countable. (Otherwise, if both the algebraic and transcendental numbers were countable, then R would be countable, contrary to fact.) Thus we see that the bulk of the irrationals is found among the real transcendental numbers.

Here we have a strange situation. There are more transcendental numbers than algebraic numbers, but in our study we have not even proved the existence of a single transcendental number. In fact, such a proof is extremely difficult and was not accomplished until the late 19th century.

The most familiar transcendental real numbers are π and e. It was not known until the 19th century that π is irrational, and not until 1882, with the proof by the German Lindeman, that it is also transcendental. There is a fascinating history of the growing understanding of π, the ratio of the circumference of a circle to its diameter. The Bible approximates it as 3; school children approximate it as $\frac{22}{7}$. For centuries it was assumed rational, and a favorite unsolved problem was that of "squaring" the circle -- finding with ruler and compass a square whose area is that of a given circle. Since operations with straight edge and compass are analogous to solutions of first - and second-degree polynomial equations, we now know that the circle cannot be "squared" because π is transcendental and, hence, cannot be the root of such an equation.

The number e, which is the lub of the set

$$\{(1 + \frac{1}{1})^1, (1 + \frac{1}{2})^2, \ldots, (1 + \frac{1}{n})^n, \ldots\},$$

is represented by 2.7182818...; it was encountered in the development of logarithms and is used as the "natural" base of logarithms. In 1873 the Frenchman Hermite proved that e is transcendental. The transcendence of real numbers such as $2^{\sqrt{2}}$ and $\log 2$ are more recent results,* known only since 1934, when it was proved that α^{β} is transcendental if α is algebraic and β is algebraic and irrational. This result establishes the transcendence of $\log r$ if r is rational and $\log r$ is irrational: By definition,

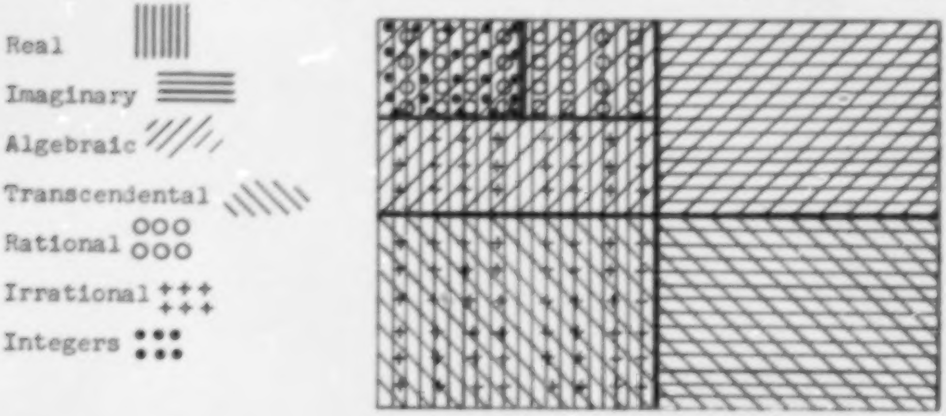
$$10^{\log r} = r.$$

---

*See Chapter 5 of the SMSG Monograph, "Rational and Irrational Numbers", by Ivan Niven, for a discussion of these results. Also see Chapter 7 for a proof of the existence of a real transcendental number.

Now if log r were algebraic and irrational, then r would be transcendental, according to the above theorem. But r is given rational; hence, log r is transcendental.

To summarize, we diagram the complex number system as follows:



(The relative areas of the regions in the above diagram do not in any way indicate the relative sizes (cardinalities) of the various sets.)

Thus, we see that every real number is either algebraic or transcendental, but not both. Every real transcendental number is irrational, but some irrational numbers are algebraic. And every rational number is algebraic, but some algebraic numbers are irrational.

Exercises

1. Prove: If A is algebraic and T is transcendental, then
   (a) A + T is transcendental,
   (b) AT is transcendental,
   (c) $\sqrt[n]{T}$ is transcendental.
2. Is the set of transcendental numbers closed under
   (a) addition, (b) multiplication, (c) division?

Answers to Exercises; pages 1.15 - 1.18:

1. $E \times O = E$, $O \times E = E$; O is an identity for ×. × is distributive through +, but + is not distributive through ×.

2. The set is closed under $\circ$ and $*$. $\circ$ is commutative, but $*$ is not. r is an identity for $\circ$. There is no identity for $*$. $\circ$ is not distributive through $*$ and $*$ is not distributive through $\circ$. Every element has an inverse under $\circ$.

3.

$\circ$	A	B	C
A	B	C	A
B	C	A	B
C	A	B	C

This set and this binary operation form an algebraic system closed under $\circ$. $\circ$ is commutative and associative. C is an identity element for $\circ$ and every element has an inverse under $\circ$. The system is a commutative group.

4. The resulting system is a commutative group.

5. Yes.

6. Yes.

7. This is not a group since v does not have an inverse and $⧺$ is not associative since, for example, $u ⧺ (u ⧺ v) = u ⧺ r = u$, but $(u ⧺ u) ⧺ v = r ⧺ v = v$.

8. Yes.

9. Yes, it is a ring. 1 is an identity for ×, but no element of I except 1 has an inverse under ×. × is commutative.

10. Yes, this is a commutative ring.

13. (a) Yes
    (b) No
    (c) Yes
    (d) No
    (e) No

Answers to Exercises; pages 2.7 - 2.9:

1. C = J, E = H, A = D.

2. C, E, G, H and J are proper subsets of W. H and E are proper subsets of G.

3. (a) one-to-one
   (b) many-to-one
   (c) one-to-one
   (d) many-to-many
   (e) many-to-one
   (f) many-to-one

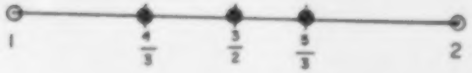
4. (a) Infinite
   (b) Not infinite
   (c) Infinite
   (d) Infinite

5. (a) Yes
   (b) No
   (c) Yes
   (d) No
   (e) No
   (f) Yes
   (g) No
   (h) Yes
   (i) No
   (j) Yes

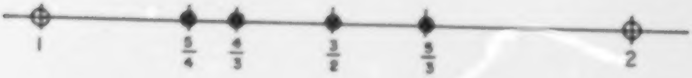
   3 is a prime but 4 is not; that is $4 = 2^2$ is the product of two numbers both greater than 1, but 3 is not and, hence, the difference in answer to (a) and (b).

   (k) No
   (l) Yes
   (m) Yes

6. (a)


1
$\frac{4}{3}$
$\frac{3}{2}$
$\frac{5}{3}$
2

(b)


1
$\frac{5}{4}$
$\frac{4}{3}$
$\frac{3}{2}$
$\frac{5}{3}$
2

Answers to Exercises; pages 2.19 - 2.21:

1. (a)

A	B	A and B	A or B	not-A	if A then B	not-A or B
T	T	T	T	F	T	T
F	T	F	T	T	T	T
T	F	F	T	F	F	F
F	F	F	F	T	T	T

(b)

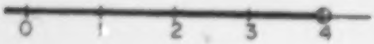
A	B	A and not-B	if not-A, then B	if B, then A
T	T	F	T	T
T	F	T	T	T
F	T	F	T	F
F	F	F	F	T

(c)

A	B	if A, then B	A or B	A and B	A and not-A
T	F	F	T	F	F
		not possible			
T	T	T	T	T	F
F	F	T	F	F	F

2. (a) {1, 2, 3}.
   (b) The set of positive integers.
   (c) ∅ = empty set.
   (d) The set of positive integers.
   (e) The set of positive integers.
   (f) The set of positive integers greater than or equal to 4. {4, 5, ...}.
   (g) {1, 2, 3, 4, ...}.
   (h) {2}.
   (i) {2}.
   (j) {1, 2}.

3. (a)



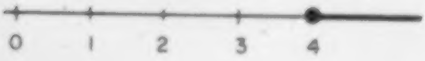
(b)

(c) ∅

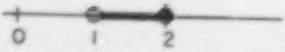
(d)

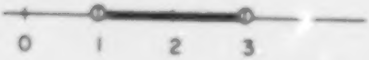
(e)

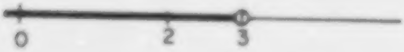
(f)



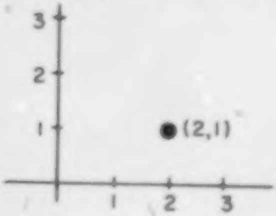
(g)

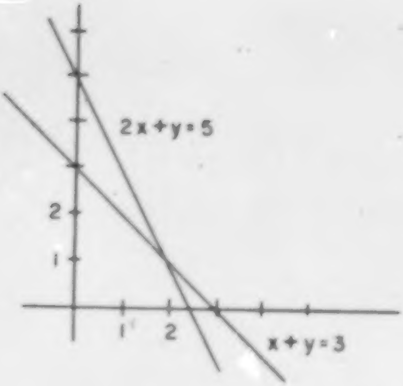
(h) 

(i) 

(j) 

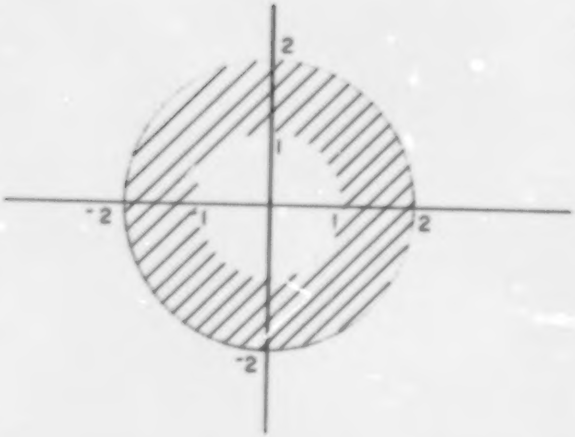
4. (a) {(1, 1), (4, 2), (9, 3), (16, 4), (25, 5), (36, 6)}.
   (b) {(1, 3)}.
   (c) {(1, 4), (2, 3), (3, 2), (4, 1), (2, 2)}.
   (d) {(1, 4)}.
   (e) {(1, 2), (1, 3), (1, 4), (1, 5), (2, 2), (2, 3)
   (2, 4), (2, 5)}.
   (f) {(-1, 1), (-2, 2), (-2, 3), (-3, 2), (-3, 3)}.

5. (a) 

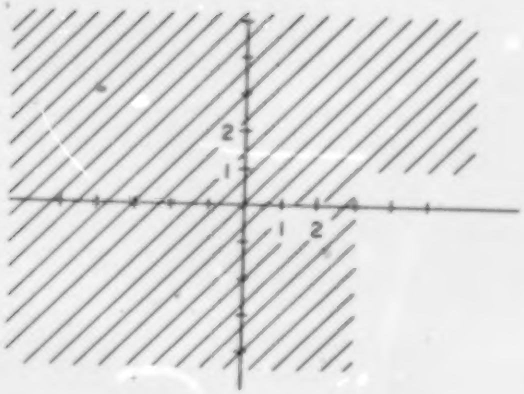
(b) 

Ans. 6

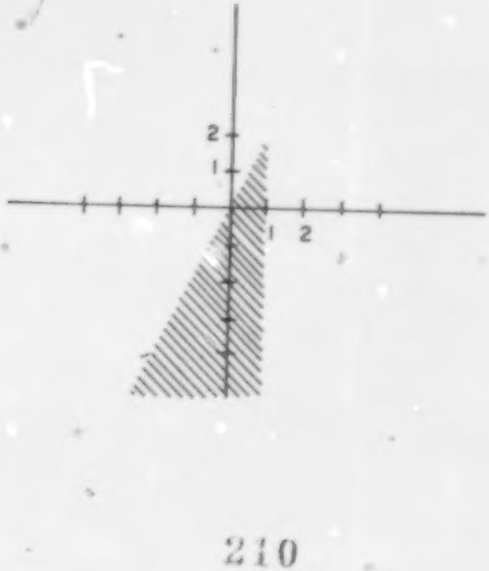
(c)



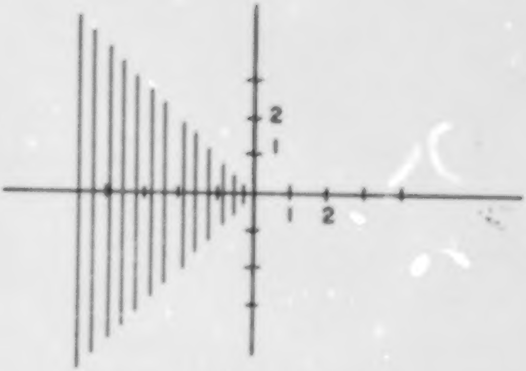
(d)



(e)



(f)



Answers to Exercises; pages 2.32 - 2.34:

1. (a)

A	B	if A, then B	not-A	not-A or B
T	T	T	F	T
F	T	T	T	T
T	F	F	F	F
F	F	T	T	T

Two statements are equivalent.

(b)

A	B	not-A	not-B	A or not-B	not-A and B
T	T	F	F	T	T
F	T	T	F	F	T
T	F	F	T	T	F
F	F	T	T	T	T

Two statements are not equivalent.

Ans. 8

(c)

A	B	if A, then B	if B, then A
T	T	T	T
F	T	T	F
T	F	F	T
F	F	T	T

Two statements are not equivalent.

(d)

A	B	not-A	not-B	if A, then B	if not-A, then not-B
T	T	F	F	T	T
F	T	T	F	T	F
T	F	F	T	F	T
F	F	T	T	T	T

Two statements are not equivalent.

(e)

A	B	not-A	not-B	A and B	not-(A and B)	not-A or not-B
T	T	F	F	T	F	F
F	T	T	F	F	T	T
T	F	F	T	F	T	T
F	F	T	T	F	T	T

Two statements are equivalent.

Ans. 9

(f)

A	B	not-A	not-B	(A or B)	not-(A or B)	not-A and not-B
T	T	F	F	T	F	F
F	T	T	F	T	F	F
T	F	F	T	T	F	F
F	F	T	T	F	T	T

Two statements are equivalent.

(g)

A	B	not-B	if A, then B	not-(if A, then B)	if A, then not-B
T	T	F	T	F	F
F	T	F	T	F	T
T	F	T	F	T	T
F	F	T	T	F	T

Two statements are not equivalent.

(h)

A	B	not-B	if A, then B	not-(if A, then B)	A and not B
T	T	F	T	F	F
F	T	F	T	F	F
T	F	T	F	T	T
F	F	T	T	F	F

Two statements are equivalent.

2. $A \Longrightarrow B \Longleftrightarrow$ not-A or B

   not-(A and B) $\Longleftrightarrow$ not-A or not-B

   not-(A or B) $\Longleftrightarrow$ not-A and not-B

   not-($A \Longrightarrow B$) $\Longleftrightarrow$ A and not-B

   The negative of a conjunction is the disjunction of the negatives; the negative of a disjunction is the conjunction of the negatives. The negative of a conditional is a conjunction.

   (a) If not-C then (not-A and not-B).

   (b) If (not-A or not-C) then not-A.

   (c) If (not-C and not-D) then (A and not-B).

   (d) If (C or D) then (not-A and B).

3. (a) Any $x \neq 0$.

   (b) Any $x \neq 0$ and $\neq 1$.

   (c) 41, 82; in fact $41 \cdot k$ for any integer k.

   (d) Any $x < 0$.

   (e) $x = 1$.

   (f) $x = 0$.

   (g) $(x = 0, y = 0)$, $(x = 1, y = 1)$. Any pair (x, x).

4. (a) Direct proof.

   (b) Proof by contradiction.

   (c) Contrapositive.

Answers to Exercises; pages 3.14 - 3.16:

8. $x = (b + c) - a$

9. $x = 1$

10. Subtraction is not commutative, $0 - 1 \neq 1 - 0$. It is not associative, since $(1 - 2) - 3 = -1 - 3 = -4$, but $1 - (2 - 3) = 1 - (-1) = 2$.

12. The set is a field with these two operations. The additive inverse of 3 is 2 and the multiplicative inverse is also 2.

13. This set is not a field, since not all non-zero elements have multiplicative inverses.

17. This is not a field since some non-zero elements, such as (0, 1), do not have multiplicative inverses.

Answers to Exercises; pages 3.24 - 3.27:

8. (a) $x > \frac{3}{2}$

   (b) $x \geq 1$

   (c) $1 < x < \frac{5}{3}$

   (d) $x > 3$ or $x < \frac{1}{2}$

   (e) $x > -\frac{1}{2}$ or $x < -3$

   (f) $x > 0$ or $x < -1$

   (g) ($x > 0$ and $x < 2$) or $x < -2$.

9. (a) $\emptyset$

   (b) $x > 1$ or $x < 0$

   (c) $x < 1$ or $x > 5$

   (d) $x = 0$ or $x = -2$

   (e) $1 < x < 3$

   (f) all x in $\mathcal{R}$ not equal to 2

   (g) $x > \sqrt{5}$ or $x < -\sqrt{5}$ or $-\sqrt{3} < x < \sqrt{3}$

   (h) $0 < x < 2$

   (i) all x in $\mathcal{R}$.

10. (a) {(0, 4), (1, 3), (-1, 3), (2, 2), (-2, 2), (3, 1), (-3, 1), (4, 0), (5, -1), (6, -2), (7, -3)}

    (b) {(-3, -1), (-3, 7), (-2, -2), (-2, 6), (-1, -3), (-1, 5), (0, 4), (1, 3), (2, 2), (3, 1), (4, 0), (5, -1), (6, -2), (7, -3)}

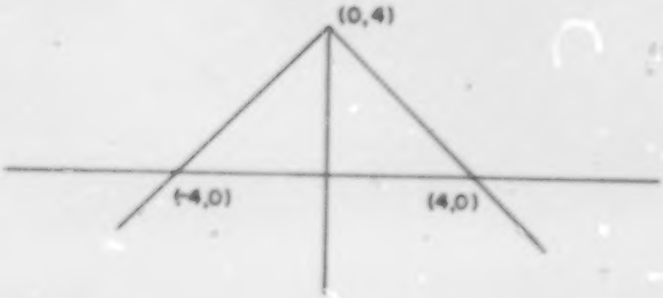
    (c) {(-1, -3), (-1, 3), (1, -3), (1, 3), (2, 2), (-2, -2), (0, 4), (4, 0), (-3, -1), (-3, 1), (3, -1), (2, -2), (3, 1), (-2, 2), (0, -4), (-4,0)}

    (d) $\emptyset$

    (e) $\emptyset$

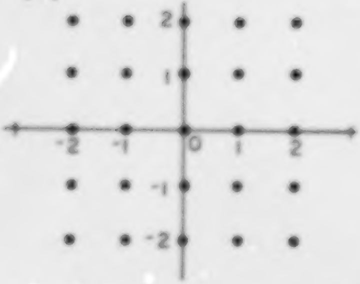
    (f) {(0, 0), (1, 0), (2, 0), (3, 0), (0, 1), (1, 1), (2, 1), (3, 1)}.

11. (a)

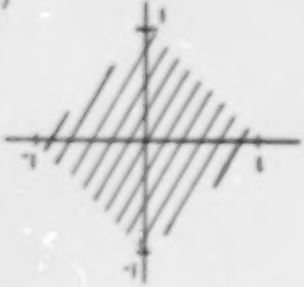


Ans. 13

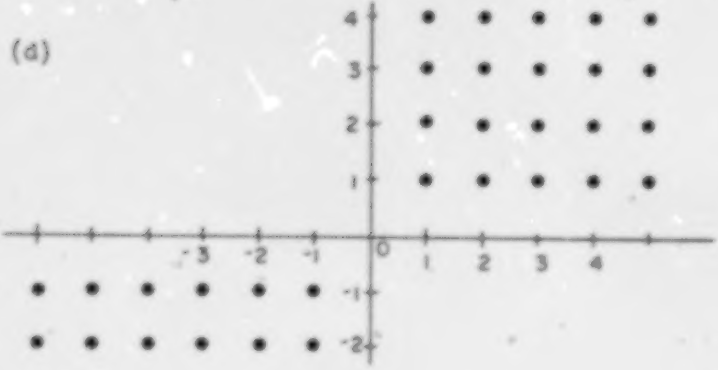
(b)



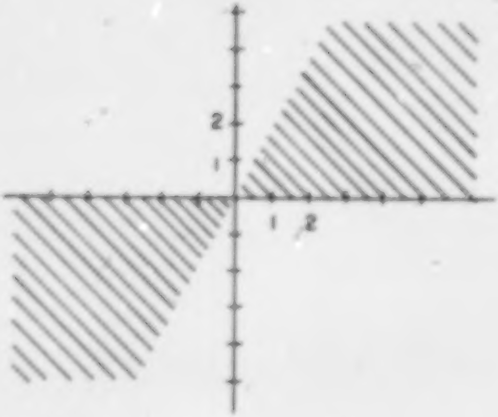
(c)



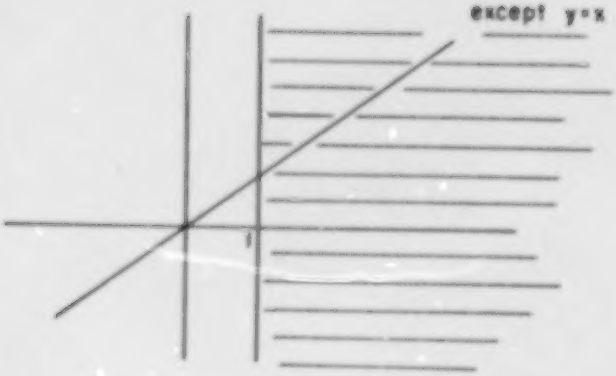
(d)


4
3
2
1
-3
-2
-1
0
1
2
3
4
-1
-2

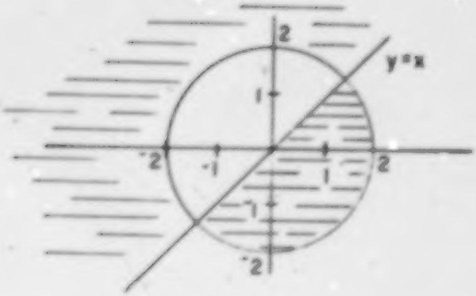
(e)



(f)



(g)



12. O1 and O2 are satisfied but O3 and O4 are not. For example, $2 < 3$ but $2 + 2 \nless 2 + 3 = 0$; $1 < 4$ but $1 + 1 \nless 1 + 4 = 0$, and $2 < 3$, $0 < 2$ but $4 \nless 2 \cdot 3 = 1$.

13. No.

Answers to Exercises; pages 3.38 - 3.39:

3. (a) $x = 3$
   (b) $x = 2$ or $x = 4$
   (c) $2 < x < 4$
   (d) $\emptyset$
   (e) $1 < x < 2$
   (f) $-5 < x \leq 3$
   (g) $x = 1$ or $x = 3$
   (h) $-3 < x < -\frac{3}{2}$

Answers to Exercises; pages 4.7 - 4.10:

2. No. There is no identity for addition; and under multiplication it is not a group since no element of N except 1 has an inverse.

10. (a) $b = 12$, $c = 2$
    (b) $b = 18$, $c = 4$
    (c) $b = 15$, $c = 10$
    (d) Impossible since if $bc = 84 = 2 \cdot 2 \cdot 3 \cdot 7$ and $b + c = 24$ we have $3|84$ and, hence, $3|b$ or $3|c$. But $3|(b + c)$ and, hence, $3|b$ and $3|c$. But then $9|bc$ which is not true.

11. (a) $(x + 2)(x + 6)$
    (b) $(x + 8)(x + 7)$
    (c) Cannot be factored in N.
    (d) $(x + 18)(x + 14)$

Answers to Exercises; pages 4.15 - 4.16:

1. (a) 0 is an even integer.

   (b) Yes.

2. (a) 90, 6, 1

3. The set of negative integers is closed under addition but not under subtraction and multiplication.

4. (a) $(932)_{ten}$ (c) $(110002)_{three}$

   (b) $(44)_{ten}$ (d) $(402)_{nine}$

5. I is a group under addition but not multiplication since 1 and -1 are the only two elements of I which have multiplicative inverses in I.

6. I is not a field because of its lack of multiplicative inverses for non-zero elements.

9. Not every composite in T can be factored uniquely into a product of primes in T. For example, 220 is in T, since $220 = 3(73) + 1$, but
   $220 = 10 \cdot 22$ and $220 = 4 \cdot 55$,
   where 10, 22, 5, 55 are primes in T.

Answers to Exercises; pages 4.31 - 4.32:

1. $-\frac{37}{61} < -\frac{12}{20} < \frac{47}{59} < \frac{4}{5}$

2. (a) $(\frac{25}{32})_{ten}$ (b) $(1.2)_{four}$

7. (a) The subset of all x with $0 < x < 1$ has no least element.

   (b) Yes. 0 is the greatest element less than every element of T.

8. $.\dot{1}\dot{4}\dot{2}\dot{8}\dot{5}\dot{7}$

9. (a) $\{1\}$

   (b) $\{1, -1\}$

   (c) $\{1, -1, \frac{3}{2}\}$

   (d) $\{1, -1, \frac{3}{2}, \sqrt{2}, -\sqrt{2}\}$

10. (a) $(x^2 - 3)(x^2 + 3)$

    (b) $(x - \sqrt{3})(x + \sqrt{3})(x^2 + 3)$

Answers to Exercises; pages 5.6 - 5.7:

1. (a) $\frac{6}{7}$, 1 are upper bounds, for example, and $\frac{5}{7}$ is the least upper bound.

   (b) -3.5, -3, 0 or any positive number are upper bounds and -3.6 is the lub.

3. A non-empty set S of real numbers is bounded below if there exists a real number M such that $s \geq M$ for every s in S. M is called a lower bound of S. A real number L is a greatest lower bound for S if:
   (1) L is a lower bound for S and
   (2) if M is any lower bound for S then $M \leq L$.

5. (a) $\frac{1}{2}$ is a lower bound and 1 is an upper bound.

   (b) 1 is a lower bound and 2 is an upper bound.

   (c) 0 is a lower bound and $\frac{1}{2}$ is an upper bound.

   (d) 1 is a lower bound and 2 is an upper bound.

6. (a) 1 is the lub. (c) 0

   (b) 1 (d) 1

Answers to Exercises; pages 5.13 - 5.14:

2. $1.414 > 1.41$ but $(1.414)^2 < 2$.

3. $1.415 < 1.42$ but $(1.415)^2 > 2$.

Answers to Exercises; pages 6.7 - 6.8:

1. (a) algebraic over R (d) algebraic over I
   (b) rational over R (e) polynomial over R
   (c) algebraic over I (f) rational over I

2. (b) $\dfrac{(\frac{\pi}{3} - 9)xy - \frac{2\pi}{3}y + 18x}{x - 2}$

   (f) $\dfrac{xy^2 - 2bxy + aby}{xy - ay - bx + ab}$

3. We know that $a - a = a + (-a) = 0$ for the real numbers and having defined $A - A = A + (-A)$ for arbitrary algebraic expressions we must define this to be 0 in order for the field properties to be satisfied. Similarly, $\frac{A}{A} = A \div A = A \cdot \frac{1}{A}$ must be 1.

4. (a) $\dfrac{4 - x^2}{x - 2} = \dfrac{(2 - x)(2 + x)}{x - 2} = \dfrac{\left((-1)(x - 2)\right)(2 + x)}{(x - 2)}$

$$= (-1)(2 + x) \cdot \frac{x - 2}{x - 2} = -(x + 2)$$

   (b) $\dfrac{x - y}{4y^2 - x^2} - \dfrac{3}{x + 2y} = \dfrac{x - y}{4y^2 - x^2} \cdot \dfrac{-1}{-1} - \dfrac{3}{x + 2y} \cdot \dfrac{x - 2y}{x - 2y}$

$$= \frac{y - x}{-4y^2 + x^2} - \frac{3x - 6y}{x^2 - 4y^2}$$

$$= \frac{y - x - 3x + 6y}{x^2 - 4y^2}$$

$$= \frac{-4x + 7y}{x^2 - 4y^2}$$

(c) $\frac{x^3 + 2}{x + 1} = \frac{x^2(x + 1) + 2 - x^2}{x + 1}$

$= \frac{x^2(x + 1) + (-x)(x + 1) + (x + 1) + 1}{x + 1}$

$= x^2 - x + 1 + \frac{1}{x + 1}$

5. (a) $x^3(x - 3)(x - 4)$ over I, F and R.

(b) $x(x - \frac{1}{2})(x - 2)$ over F and R.

(c) $(y + \sqrt{2} - a)(y - \sqrt{2} - a)$ over R.

(d) $(a^2 - 2a + 2)(a^2 + 2a + 2)$ over I, F, R.

Answers to Exercises; pages 6.17 - 6.18:

1. (a) $\{-3\}$

(b) $\{-3, \frac{1}{2}\}$

(c) $\{-3, \frac{1}{2}, \sqrt{3}, -\sqrt{3}\}$

(d) $\{1, 2\}$

(e) $\{(1, 1), (2, 1)\}$

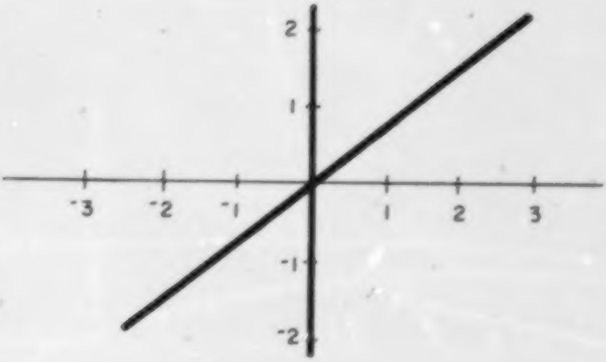
(f) $\{(1, 1), (2, 1)\}$

(g) The truth set is empty.

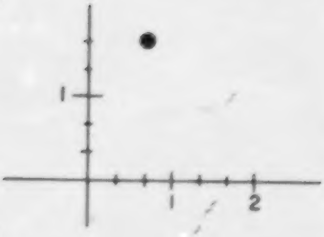
(h) The truth set is F.

(i) All x in R with $x > 1$.

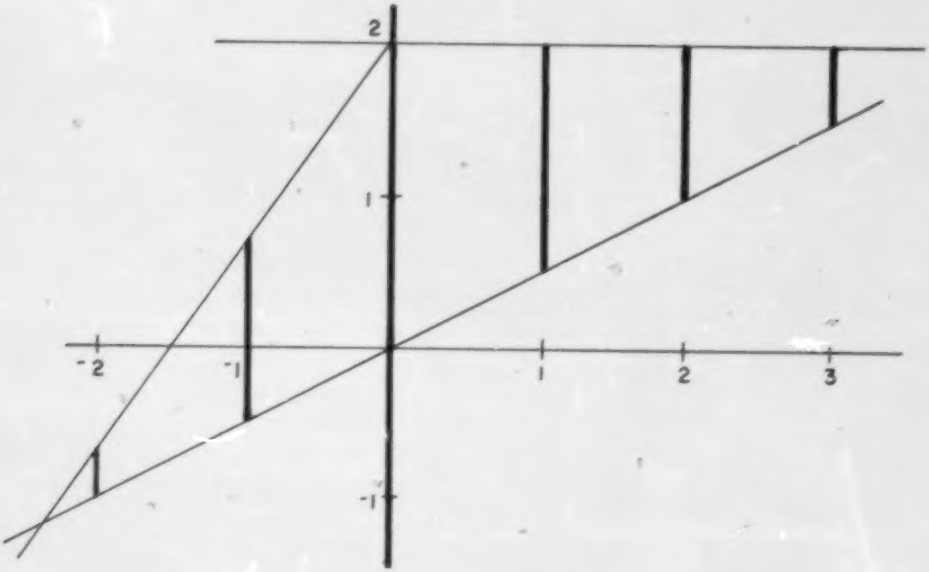
2. (a)



(b)

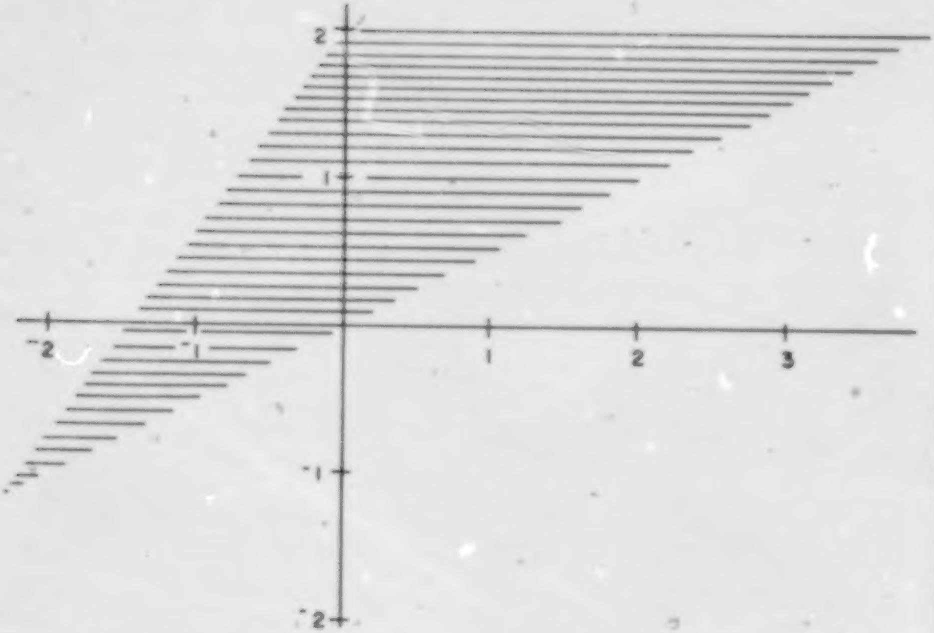


(c)

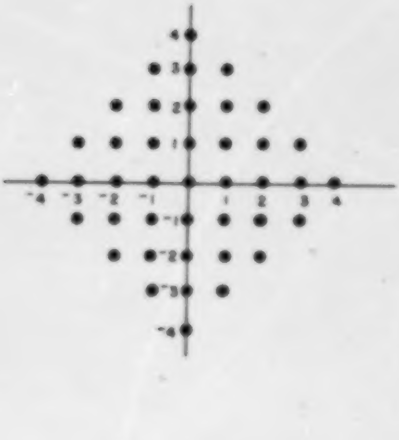


Ans. 22

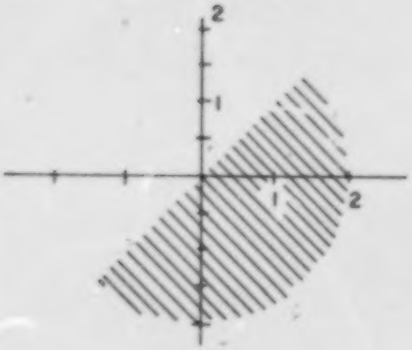
(d)



(e)



(f)



3. (a) $\{1, \frac{2}{5}\}$

(b) $\{1, 2, -1\}$

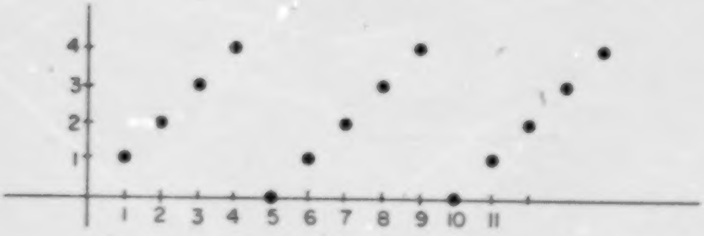
(c) $\emptyset$

(d) $\{0, 1\}$

(e) $\{\frac{1 + \sqrt{57}}{2}, \frac{1 - \sqrt{57}}{2}\}$

Answers to Exercises; pages 6.24 - 6.26:

1. (a) The domain is the set of positive integers and the range is the set $\{0, 1, 2, 3, 4\}$.
$\{(1, 1), (2, 2), (3, 3), (4, 4), (5, 0), (6, 1), (7, 2), (8, 3), (9, 4), \ldots\}$. As a graph

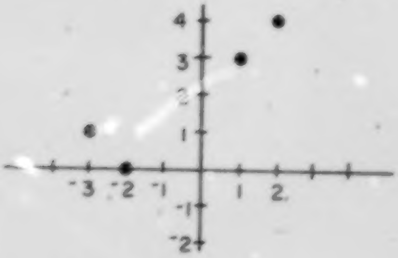


$f: n \longrightarrow r_o$ where $n = 5 \cdot k + r_o$, $0 \leq r_o < 5$.

(b) The domain is the set of natural numbers and the range is the set of natural numbers of the form $3n + 2$, n a natural number. $f(n) = 3n + 2$.

To each natural number n there is assigned the number obtained by adding 2 to 3 times n. $f{:}n \longrightarrow 5 + 3(n - 1)$.

(c) The domain is the set $\{-3, -2, 1, 2\}$ and the range is the set $\{1, 0, 3, 4\}$.
$\{(-3, 1), (-2, 0), (1, 3), (2, 4)\}$.



$f(x) = \sqrt{(x + 2)^2}$

(d) The domain is the set of all real numbers x such that $0 \leq x \leq 3$; the range is the set of real numbers y such that $0 \leq y \leq 2$.

$f{:}x \longrightarrow -\frac{2}{3}x + 2$.

$y = -\frac{2}{3}x + 2$.

To each x in R between 0 and 3 there is assigned the real number y which is equal to $-\frac{2}{3}$ times x plus 2.

Ans. 25

2. (a) All non-zero real numbers.

(b) All real numbers less than or equal to -2 or greater than or equal to 2.

(c) All real numbers x such that $x > 1$ or $x \leq 0$.

(d) R.

3. (a) The domain of F is contained in the domain of f and for x in the domain of F, $(x \neq -2)$, $f(x) = F(x)$.

(b) g and G define the same function.

(c) h and H define the same function.

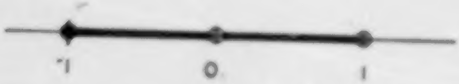
4. (a) $f(-\frac{1}{2}) = -1$, $f(\sqrt{5})$ not defined, $f(\frac{3}{2}) = \frac{3}{2}$.

(b) The set of all x in R with $-1 \leq x < 0$ or $0 < x \leq 2$.

(c) The range of f is the set containing -1 and all real numbers less than or equal to 2 and greater than zero.

(d) The set of all x in R with $0 < x \leq 2$ or $x = -1$.

(e)



5. $g(-t) = t^2 - 1 = g(t)$

$-g(t) = 1 - t^2$

$2g(t) = 2(t^2 - 1)$

$g(2t) = 4t^2 - 1$

$g(t - 1) = t^2 - 2t$

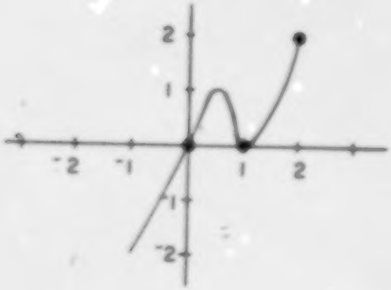
$g(t) - 1 = t^2 - 2$

$g\left(g(t)\right) = t^4 - 2t^2$

$g\left(g(\frac{1}{t})\right) = \frac{1 - 2t^2}{t^4}$

$g(\frac{1}{g(t)}) = \frac{2t^2 - t^4}{(t^2 - 1)^2}$

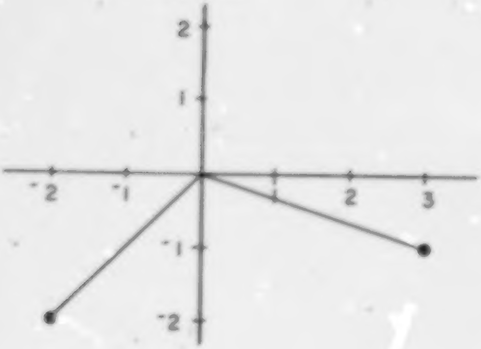
6.



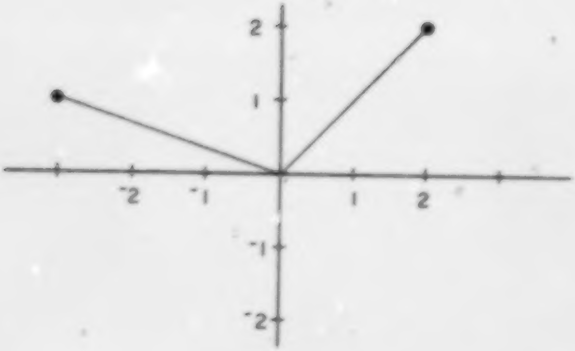
This is the graph of one function satisfying the conditions. There are infinitely many such functions.

7. (a), (c), and (f) define functions.

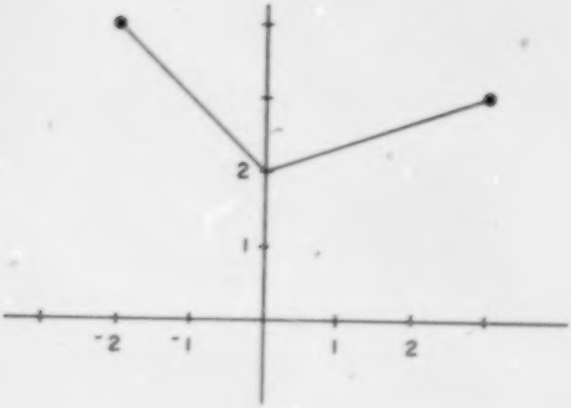
8. (a)



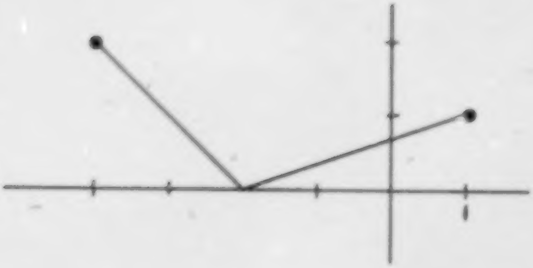
(b)



(c)



(d)



Answers to Exercises; pages A.7 - A.8:

1. (a) $\frac{41}{333}$ (d) $\frac{1}{7}$

   (b) $\frac{194}{45}$ (e) $\frac{635}{100}$

   (c) $\frac{1}{275}$ (f) $\frac{11}{100}$

2. (a) {1, 1.7, 1.73, 1.732}

   (b) {.3, .33, .333, .3333}

   (c) {1, 1.2, 1.25, 1.259}

   (d) {2, 2.2, 2.23, 2.236}

3. (a) For example, 2.3, 2.31, 2.34, ..., 2.4 or any terminating decimal between 2.3 and 2.4.

   (b) 6.6, 6.61, 6.624973, ..., 6.63 or any terminating decimal between 6.6 and 6.63.

   (c) Any terminating decimal larger than $.\dot{9}\dot{3}$ and less than $.\dot{9}\dot{6}$. For example, .9404, .95, .95999, .96.

5. $3.1416 = \frac{31416}{10000}$ and $\frac{22}{7}$ are rational numbers and π is not.

6. (a) 1 (c) $\frac{101}{110}$

(b) $\frac{1}{11}$ (d) $\frac{10011}{1110}$

Answers to Exercises; pages B.10 - B.11:

3. (a) $\{2i, -2i\}$

(b) $\{-\frac{1}{2} + \frac{\sqrt{3}}{2}i, -\frac{1}{2} - \frac{\sqrt{3}}{2}i\}$

(c) $\{0, 1 + \frac{\sqrt{10}}{2}, 1 - \frac{\sqrt{10}}{2}\}$

(d) $\{3, \sqrt{5}, -\sqrt{5}, 3i, -3i\}$

4. (a) $(-1, 1)$ (e) $(10, 0)$

(b) $(\frac{2}{3}, \frac{4}{3})$ (f) $(-7, -19)$

(c) $(-7, 19)$ (g) $(-4, 10)$

(d) $(-\frac{37}{10}, \frac{11}{10})$ (h) $(6, 42)$

Answers to Exercises; page C.6:

2. The set of transcendental numbers is not closed under addition, multiplication or division.
